PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/73, 15/83, 15/82
C12P 41/00, C12N 15/53, 15/55
C12N 5/10, 7/01 // C12N 9/00

A1

(11) International Publication Number: **WO 90/00611**

(43) International Publication Date: 25 January 1990 (25.01.90)

(21) International Application Number: PCT/US89/03060

(22) International Filing Date: 14 July 1989 (14.07.89)

(30) Priority data:
219,279 15 July 1988 (15.07.88) US
363,138 8 June 1989 (08.06.89) US

(71) Applicant: BIOSOURCE GENETICS CORPORATION [US/US]; 3333 Vaca Valley, Vacaville, CA 95688 (US).

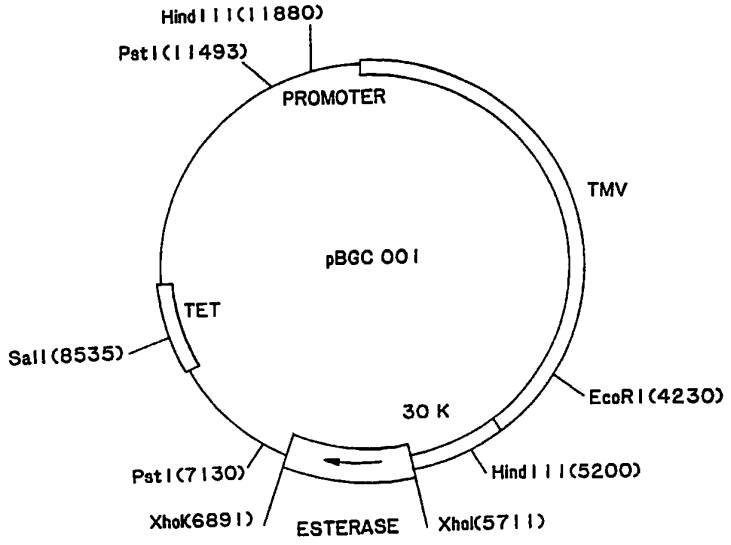
(72) Inventors: ERWIN, Robert, L. ; 417 Arkansas Street, San Fracisco, CA 94022 (US). GRILL, Laurence, K. ; 2077 Jackson Street, ,101, San Francisco, CA 94109 (US).

(74) Agents: IHNEN, Jeffrey, L. et al.; Robbins & Laramie, 2100 Pennsylvania Avenue, N.W., Suite 600, Washington, DC 20037 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: SYNTHESIS OF STEREOSPECIFIC ENZYME BY NON-CHROMOSOMAL TRANSFORMATION OF A HOST



(57) Abstract

The present invention relates to the synthesis of a stereospecific enzyme by the non-chromosomal transformation of a host. The transformation is accomplished by a viral vector which is non-infective but which is self-replicating and which contains a coding sequence for an enzyme which is capable of enantiospecific catalysis of an organic chemical.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CM	Cameroon
DE	Germany, Federal Republic of
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
HU	Hungary
IT	Italy
JP	Japan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
MC	Monaco
MG	Madagascar
ML	Mali
MR	Mauritania
MW	Malawi
NL	Netherlands
NO	Norway
RO	Romania
SD	Sudan
SE	Sweden
SN	Senegal
SU	Soviet Union
TD	Chad
TG	Togo
US	United States of America

TITLE OF THE INVENTION

SYNTHESIS OF STEREOSPECIFIC ENZYME BY NON-CHROMOSOMAL TRANSFORMATION OF A HOST

BACKGROUND OF THE INVENTION

The present invention relates to the synthesis of compounds by the non-chromosomal transformation of a host. The transformation is accomplished by a viral vector which is non-infective but self-replicating and which contains a coding sequence for a stereoselective enzyme. More particularly the present invention also relates to the purification of a racemic mixture by stereospecific enzymes. Purification can be accomplished by selectively converting a precursor compound from a racemic mixture of that compound to a desired stereoisomer. Purification can be accomplished by selectively converting one member of a racemic mixture to another compound, thereby leaving the desired stereoisomer which can then be purified.

A. VIRUSES

Viruses are a unique class of infectious agents whose distinctive features are their simple organization and their mechanism of replication. In fact, a complete viral particle, or virion, may be regarded mainly as a block of genetic material (either DNA or RNA) capable of autonomous replication, surrounded by a protein coat and sometimes by an additional membranous envelope such as in the case of alpha viruses. The coat protects the virus from the environment and serves as a vehicle for transmission from one host cell to another.

Unlike cells, viruses do not grow in size and then divide, because they contain within their coats few or none of the biosynthetic enzymes and other machinery required for their replication. Rather, viruses

multiply in cells by synthesis of their separate components, followed by assembly. Thus the viral nucleic acid, after shedding its coat, comes into contact with the appropriate cell machinery where it specifies the synthesis of proteins required for viral reproduction. The viral nucleic acid is then itself replicated through the use of both viral and cellular enzymes. The components of the viral coat are formed and the nucleic acid and coat components are finally assembled. With some viruses, replication is initiated by enzymes present in virions.

Viruses are subdivided into three main classes -- animal viruses, plant viruses and bacterial viruses. Within each class, each virus is able to infect only certain species of cells. With animal and bacterial viruses, the host range is determined by the specificity of attachment to the cells which depends on properties of both the virion's coat and specific receptors on the cell surface. These limitations disappear when transfection occurs, i.e., when infection is carried out by the naked viral nucleic acid, whose entry does not depend on virus-specific receptors.

A given virus may contain either DNA or RNA, which may be either single- or double-stranded. The portion of nucleic acid in a virion varies from about 1% to about 50%. The amount of genetic information per virion varies from about 3 to 300 kb per strand. The diversity of virus-specific proteins varies accordingly. Examples of double-stranded DNA containing viruses include, but are not limited to, Hepatitis B virus, papovaviruses such as polyoma and papilloma, adenovirus, poxviruses such as vaccinia, caulimoviruses such as Cauliflower mosaic virus (CaMV), *Pseudomonas* phage PMS2, Herpesvirus, *Bacillus subtilis* phage SP8, and the T bacteriophages. Representative viruses which are single-stranded DNA are the parvoviruses and the

bacteriophages ϕX174, f1 and M13. Reoviruses, cytoplasmic polyhedrosis virus of silkworm, rice dwarf virus and wound tumor virus are examples of double-stranded RNA viruses. Single-stranded RNA viruses include tobacco mosaic virus (TMV), turnip yellow mosaic virus (TYMV), picornaviruses, myxoviruses, paramyxoviruses and rhabdoviruses. The RNA in single-stranded RNA viruses may be either a plus (+) or a minus (-) strand. For general information concerning viruses see Grierson, D. et al., *Plant Molecular Biology*, Blackie, London, pp. 126-146 (1984); Dulbecco, R. et al., *Virology*, Harper & Row, Philadelphia (1980); White, A. et al., *Principles of Biochemistry*, 6th Ed., McGraw-Hill, New York, pp. 882-900 (1978).

One means for classifying plant viruses is based on the genome organization. Although many plant viruses have RNA genomes, the organization of genetic information differs between groups. The genome of most monopartite plant RNA viruses is a single-stranded molecule of (+)- sense. There are at least 11 major groups of viruses with this type of genome. An example of this type of virus is TMV. At least six major groups of plant RNA viruses have a bipartite genome. In these, the genome usually consists of two distinct (+)-sense single-stranded RNA molecules that are encapsidated in separate particles. Both RNAs are required for infectivity. Cowpea mosaic virus (CPMW) is an example of a bipartite plant virus. The third major type, containing at least six major types of plant viruses, has three (+)-sense single-stranded RNA molecules, i.e., is tripartite. Each strand is separately encapsidated and all three are required for infectivity. An example of a tripartite plant virus is alfalfa mosaic virus (AMV). Many plant viruses also have smaller sub-genomic mRNAs that are synthesized to amplify a specific gene product. One group of plant viruses which have a

single-stranded DNA genome are the geminiviruses, such as cassava latent virus and maize streak virus.

Techniques have been developed which are utilized to transform many species of organisms. Hosts which are capable of being transformed by these techniques include bacteria, yeast, fungus, animal cells and plant cells or tissue. Transformation is accomplished by using a vector which is self-replicating and which is compatible with the desired host. The vectors are generally based on either a plasmid or a virus. Foreign DNA is inserted into the vector, which is then used to transform the appropriate host. The transformed host is then identified by selection or screening. For further information concerning the transformation of these hosts, see Maniatis, T. et al., *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor (1982); *DNA Cloning*, Ed. Glover, D.M., IRL Press, Oxford (1985); Grierson, D. et al., *supra*; and *Methods in Enzymology*, volumes 68, 100, 101, 118 and 152-155 (1979, 1983, 1983, 1986 and 1987).

Viruses that have been shown to be useful for the transformation of appropriate hosts include bacteriophages, animal viruses such as adenovirus type 2 and vaccinia virus and plant viruses such as CaMV and brome mosaic virus (BMV). An example of the use of a bacteriophage vector is shown in U.S. Patent 4,508,826. U.S. Patent 4,593,002 shows the use of adenovirus type 2 as well as a bacteriophage for the transformation of the appropriate host. The use of a vaccinia virus is shown in U.S. Patent 4,603,112. Transformation of plants using plant viruses is described in EP A 67,553 (TMV), EP A 194,809 (BMV) and Brisson, N. et al., *Methods in Enzymology 118*, 659 (1986) (CaMV).

When the virus is a DNA virus, the constructions can be made to the virus itself. Alternatively, the virus can first be cloned into a bacterial plasmid for

eas of constructing the desired viral vector with th foreign DNA. The virus can then be excised from the plasmid. If the virus is a DNA virus, a bacterial origin of replication can be attached to the viral DNA which is then replicated by the bacteria. Transcription and translation of this DNA will produce the coat protein which will encapsidate the viral DNA. If the virus is an RNA virus, the virus is generally cloned as a cDNA and inserted into a plasmid. The plasmid is then used to make all of the constructions. The RNA virus is then produced by transcribing the viral sequence of the plasmid and translation of the viral genes to produce the coat protein(s) which encapsidate the viral RNA.

B. STEREOISOMERS AND ENZYMATIC SEPARATION

It is known that many biologically active compounds exist as a mixture of stereoisomers, especially when these compounds are chemically synthesized. Up to now these mixtures are frequently used as such in agricultural and pharmaceutical applications. Usually the desired biological activity resides in one stereoisomer so that in case of a two-stereoisomer mixture the potency of the mixture is reduced by half. Still a major reason for the use of mixtures of stereoisomers is that the cost of separation of the stereoisomers exceeds the potential advantage of a possible increase in activity. However, it is apparent that modern pharmacologists are becoming increasingly aware of other implications of the administration of mixtures wherein one or more stereoisomers have to be regarded as an impurity that may not have the desired therapeutic effect, but may even have other unwanted physiological effects including toxicity.

A number of α-methylarylacetic acids (2-arylpropionic acids) are known as anti-inflammatory agents. Among the best known are ibuprofen, flurbiprofen, keto-

profen and suprofen (all of which are substituted α-methylbenzeneacetic acids), and naproxen (a substituted α-methylaphthaleneacetic acid). As is well known, the α-methylarylacetic acid molecule is chiral at the α-carbon atom and therefore exists in two stereoisomeric forms, the *R*- and *S*-forms (these forms are named by application of the "Sequence Rule," see *J. Org. Chem. 35*, 2863 (1970)). The *S*-enantiomers of these α-methylarylacetic acids generally possess greater anti-inflammatory activity than the *R*-enantiomers ("Non-Steroidal Antiinflammatory Drug," J.G. Lombardino Ed., John Wiley & Sons, New York, 1985, p. 303). More particularly, it has been discovered that the *in vitro* anti-inflammatory activity of naproxen as well as ibuprofen resides in the *S*-enantiomer (optically active stereoisomer), which is up to 150 times as active as its antipode, as known. e.g., from Adams, S. et al., *J.Pharm.Pharmac. 28*, 256 (1976) and Hutt, A.J. et al., *Clin.Pharmacokin. 9*, 371 (1984).

More specifically, the pharmaceutically active compound produced in a stereospecific *S*-configuration of formula (I) or a pharmaceutically acceptable salt or ester thereof, preferably by an alkali metal salt or an alkaline earth metal salt thereof, is one wherein R_1 is an aryl group. The aryl group is preferably a monocyclic, polycyclic, or condensed polycyclic aromatic or heteroaromatic group having up to 12 (preferably 6 to 12) carbon atoms in the aromatic system such as phenyl, biphenyl, naphthyl, thienyl and pyrrolyl. The aromatic group is optionally substituted with one or more nitro, halo, hydroxy, C_{1-4} alkyl, C_{3-6} cycloalkyl, benzyl, C_{1-4} alkoxy, C_{1-4} alkylthio, C_{1-4} haloalkyl, C_{1-4} haloaolkoxy, phenoxy, thenoyl and benzoyl groups.

Specific examples of aryl groups suitable for the purposes of the present invention are phenyl, 4-benzoyl-phenyl, 4-isobutylphenyl, 40(2-thienoyl)phenyl, 3-

fluorobiphenyl, 6-methoxy-2-naphthyl, 5-halo-6-methoxy-2-naphthyl, 6-hydroxy-2-naphthyl and 5-halo-6-hydroxy-2-naphthyl.

The pharmaceutically active compound (I) is produced by the stereospecific hydrolysis of an ester of

(II)

wherein R_1 is as defined above and R_2 is preferably a straight, branched or cyclic alkyl group having from 1 to 12 carbon atoms, optionally substituted with phenyl or one or more electron-withdrawing substituents, for example, halo, nitro, cyano, hydroxy, C_{1-4} alkoxy, C_{1-4} alkylthio or $-C(O)R_3$ wherein R_3 is C_{1-4} alkyl, C_{3-6} cycloalkyl, hydroxy, C_{1-4} alkoxy, C_{3-6} cycloalkoxy, phenoxy, benzyloxy, NR_1R_5 (in which R_1 and R_5 are independently H, C_{1-4} alkyl, C_{3-6} cycloalkyl, or jointly form a 5- or 6-membered ring together with the nitrogen, the ring optionally including a hetero group selected from O, NH or N-(C_{1-4} alkyl)), or -OM wherein M is an alkali metal.

The electron-withdrawing substituents, if present, are preferably at the α- or ß-position of the R_2 group, to an extent consistent with the stability of the group. Esters in which the R_2 groups contain electron-withdrawing substituents are referred to as activated ester, since they generally hydrolyze more rapidly than those where the R group is not so substituted.

Specific examples of R_2 alkyl groups are methyl, ethyl, butyl, hexyl, octyl, dodecyl, benzyl, 2-chloroethyl, 2,2,2-trichloroethyl, 2-fluoroethyl, 2,2,2-trifluoroethyl, 2-bromoethyl, cyanomethyl, 2-nitropropyl, carboethoxymethyl, methoxymethyl, 2-

hydroxy-1,2-dimethoxy-carbonylethyl, 2-hydroxy-1,2-dicarboxyethyl, 2-hydroxy-1,2-diethoxycarbonylethyl, etc.

If the resulting (S)-α-methylarylacetic acid is a precursor of a drug such as naproxen, for example (*S*)-5-halo-6-methoxy-α-methyl-2-naphthaleneacetic acid, (*S*)-6-hydroxy-α-methyl-2-naphthaleneaceetic acid of (*S*)-5-halo-6-hydroxy-α-methyl-2-naphthaleneacetic acid, such precursor can be converted to naproxen by methods described in EPA 95 901 (1983).

Irichijima, S. et al., *Agric. Biol. Chem. 45*, 1389 (1981) showed that selected microorganisms are able to hydrolyze the methyl esters of naproxen and ketoprofen but with low conversions. *Aspergillus sojae* preferentially hydrolyzed the methyl ester of *R*-naproxen to produce naproxen having 95% by weight of the *R*-configuration while *Mycobacterium smegmatis* preferentially hydrolyzed the methyl ester of *S*-ketoprofen to give ketoprofen having only 69% by weight of the *S*-configuration.

The production of *S*-naproxen from racemic naproxen esters is described in German application DE 3 345 660. The *S*-naproxen is not directly formed from the racemic mixture of naproxen esters, but is formed by saponification or by enzymatic hydrolysis of the *S*-naproxen ester, which remains in the reaction mixture after the ester of *R*-naproxen is enzymatically hydrolyzed by a microbial enzyme and the *R*-naproxen formed is separated from the ester of *S*-naproxen.

EPA 0 233 656 describes an improved selective synthesis for the preparation of particularly the *S*-stereoisomer of compound (I), which comprises subjecting the ester compound (II) to stereospecific hydrolysis using an esterase produced by or derived from a microorganism. The esterase has the ability for stereoselective hydrolysis of compound (II) into

compound (I), having at least 80% by weight of the *S*-configuration, and if desired, for converting compound (I) into a pharmacologically acceptable salt or ester thereof.

Preferably, compound (I) is naproxen, ibuprofen, suprofen, fenoprofen, ketoprofen, benoxaprofen, carprofen, cicioprofen, pirprofen, lisiprofenum, flurbiprofen, fluprofen, clidanac, tertprofen, hexaprofen, indoprofen, mexoprofen, pranoprofen, R 803, protizinic acid, tiaprofenic acid or brofezil.

EPA 0 227 078 discloses a similar procedure in which a lipase is utilized to prepare the desired stereoisomer.

Examples of stereosepcific enzymes that separate racemic mixtures are not limited to those which are useful in preparing α-methylacetic acids. Pig liver esterase has been reported to catalyze the enantiospecific hydrolysis of *meso*-dimethyl cyclopropane-, cyclobutane-, and cyclohexane-1,2-dicarboxylates, giving acid-ester products that are readily convertible into 2-lactones of >97% enantiomeric excess (ee). Sabbioni, G. et al., *J. Org. Chem. 52(20)*, 4565 (1987). Pig liver esterase has also been reported to catalyze the enantiospecific hydrolysis of dimethyl 3-hydroxy-, 3-methoxyethoxymethoxy-, and 3-benzyloxyglutarates to give the corresponding acid-ester products having a 16-40% enantiomeric excess. Lister, K. et al. *Can. J. Chem. 66*, 1422 (1988). The same team of scientists which reported the characteristics of pig liver esterase also found enantiomeric selectivity in horse liver alcohol dehydrogenase. This enzyme was found to catalyze reductions of racemic cis and trans bicyclic O- and S-heterocyclic ketones with high enantiomeric selectivity. Lister, K. et al., *J. Org. Chem. 53*, 1611 (1988).

U.S. Patent No. 4,659,671 to Klibamov, A.M. discloses a method of resolving a racemic mixture of hydroxy compounds to obtain the biologically active optical isomer of a variety of chemicals, food additives and drugs. A racemic mixture of a hydroxy compound is phosphorylated to form the D, L monophosphate esters. The mixture is then treated with a stereospecific phosphatase enzyme which causes hydrolysis of one of the optical isomers. The resulting hydrolyzed isomer is then separated from the remaining monoorthophosphate ester. Klibamov contemplated the resolution of primary, secondary and tertiary alcohols by enzymes either in a liquid medium or immobilized on a solid support. Some sources of phosphatase enzyme are wheat germ, potato, calf intestine and *Escherichia coli*.

Hydroxy compounds have a wide variety of practical uses as pharmaceuticals, flavorings, agricultural chemicals and food additives. Some of the better known hydroxy compounds are threonine, malic acid, tartaric acid, menthol, carnitine and the drugs, ephedrine, octopamine, epinephrine and phenylephrine.

U.S. Patent No. 4,745,066 and European Patent Application No. 0 189 878 to Hamaguchi, S. et al. discloses a method of resolving racemic mixtures of a compound to produce an optically active glycol derivative. Sixteen (16) different enzymes are listed. Most are lipases derived from swine and a variety of microorganisms which will produce optically active glycols having greater than 99% enantiomeric excess. Many of the products derived from this method are versatile starting materials for the production of optically active pharmaceutical products and agricultural chemicals.

European Patent Application 0 258 666 to Davis, P. et al. describes a method of making chrysanthemic, lavandulic and analogous acids from racemic mixtures of

the corresponding alcohols. Davis and co-workers found that when certain microorganisms of the genus *Aspergillus* were contacted with one of the above mentioned alcohols, the alcohol was transformed stereoselectively to the corresponding acid. The stereoisomer (1R, 3R) (+)- *trans* - chrysanthemic acid was found to have superior insecticidal activity.

U.S. Patent No. 4,588,694 to Hamaguchi, S et al. describes a process of preparing an optically active oxazolidinone derivative [(S)-I] by utilizing enzymes having stereospecific esterase activity or by using a microorganism containing the enzyme. The stereospecific enzyme asymetrically hydrolyzes the racemates of the acyloxyoxazolidinone derivative [(R, S)-II]. The unreacted compound [(S)-II] is then separated from the hydrolyzed compound [(R)-I] and then hydrolyzed. The compounds [(S)-II] and [(S)-I] are important intermediates for preparing optically active β-adrenergic blocking agents (β-blockers).

Microorganisms such as *Enterobacter*, *Klebsiella*, *Micrococcus*, *Escherichia*, *Pseudomones*, *Staphylococcus*, *Alcaligenes*, and *Achromobacter* were found to produce enzymes that would perform the desired stereospecific hydrolysis. The [(R,S)-II] can be added directly to the culture medium or to the extracted enzyme to produce the desired result.

International Patent Application PCT/US86/01282 discloses a method of preparing optically active 4-hydroxy cyclopent-2-en-1-one and 4-hydroxy cyclopentan-1-one by incorporating a racemic mixture of a 3,5-dihydroxycyclopent-1-ene into a microbial growth medium. The products have ready applicaiton in the synthesis of prostoglandins. The inventors list several microorganisms which contain the necessary stereospecific enzymatic activity. Among them are

Escherichia, Mycobacterium, Bacillus, Arthrobacter, Saccharomyces and *Penicillium*.

Similar procedures have been widely published for the optical resolution of amino acids. U.S. Patent No. 4,016,037 to Mitsugi et al. describes a method of producing L-amino acids by incorporating hydantoin compounds with the growth medium of *Flavobacterium aminogenes* and allowing the L-form of the amino acid to accumulate.

European Patent Application No. 0 175 312 discloses a method of preparing optically active N-carbamoyl-α-amino acids from a racemic mixture. The racemic mixture is placed in the growth medium of a microorganism such as *Aerobacter, Bacillus, Corynebacterium* or *Nocardia*. The stereospecific enzymatic activities of the microorganism converts one member of the racemic mixture into an optically active hydantoin leaving the other member in its optically active state.

U.S. Patent No. 4,670,395 to Kung et al. discloses the stereospecific hydrolysis of the L-isomer in a mixture of L and D α-N-acyl-α-amino acid ester where the alpha carbon atom is chiral, to yield a mixture containing the D-α-N-acyl-α-amino acid ester and a L-α-amino acid. The process is performed in the presence of a combination of esterase and acylase enzymes.

SUMMARY OF THE INVENTION

The present invention relates to viral vectors which are biologically contained, self-replicating and capable of the non-nuclear chromosomal transformation of a host. The viral vectors may contain a heterologous coding sequence for an enzyme adjacent a viral promoter. The invention further relates to viruses containing the viral vectors which are transmissible, i.e. infective. A host is infected by the viruses of the invention.

Growth of the host results in the production of the enzyme which is useful for the preparation of a pharmaceutically active compound of formula (I) and other commercially useful compounds. Production cells are disclosed which are capable of producing the viruses or parts thereof.

The invention also relates to production cells which are capable of producing the viruses or parts thereof. Host cells are infected by the viruses of the invention. A process for the production of a desired product by growing the infected hosts is also within the scope of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a restriction map of pNAPT-2 in which the shaded area contains an esterase gene.

Figure 2 illustrates a restriction map of pNAPT-7 in which the shaded area contains an esterase gene.

Figure 3 illustrates a restriction map of pNAPT-8 in which the shaded area contains an esterase gene.

Figure 4 shows the 5' end of the *Hpa*II-*Bam*HI fragment of pAT63.

Figure 5 illustrates a map of pBGC 001 containing an esterase gene.

DETAILED DESCRIPTION OF THE INVENTION

The present invention includes (a) viral vectors which are non-infective but which are self-replicating and capable of the non-nuclear chromosomal transformation of a host, (b) viral vectors which contain a heterologous coding sequence adjacent a viral promoter, (c) viruses containing the viral vectors which are infective, (d) production cells which are capable of producing the viruses or parts thereof, (e) a host

infected by the viruses of the invention, and (f) a process for the production of a desired product by growing the infected hosts. The coding sequence codes for an enzyme which is capable of purifying a racemic mixture to obtain a desired stereoisomer.

In order to provide a clear and consistent understanding of the specification and the claims, including the scope given to such terms, the following definitions are provided:

Adjacent: A position in a nucleotide sequence immediately 5' or 3' to a defined sequence.

Animal Tissue: Any tissue of an animal in the organism or in culture. This term is intended to include a whole animal, animal cell, animal organ, protoplast, cell culture or any group of animal cells organized into a structural and functional unit.

Anti-sense Mechanism: A type of gene regulation based on controlling the rate of translation of mRNA to protein due to the presence in a cell of an RNA molecule complementary to at least a portion of the mRNA being translated.

Biologically Contained: The viral nucleic acid is not capable of naturally infecting a host since it is not capable of expressing a biologically functional coat protein.

Biologically Functional: The capability of performing an expected biological function in a cell or organism. For example, the biological function of a viral coat protein is the encapsidation of the viral nucleic acid. A non-biologically functional coat protein is not capable of encapsidating viral nucleic acid. A nucleotide sequence which lacks a biologically functional protein coding sequence will produce either no protein or will produce a protein which will not perform an expected function. If the entire coding sequence for the protein is removed, then no protein

will be produc d. If a significant portion of the coding sequence for the protein is removed, then any protein that is produced will not function as the entire protein would function. If the coding sequence for the protein is mutated such as by a point mutation, the protein might not function as a normal protein would function. For example, a nucleotide sequence which lacks a biologically functional coat protein coding sequence is a nucleotide sequence which does not code for a coat protein capable of encapsidating viral nucleic acid. This term is intended to include a complete deletion of the coat protein sequence.

Cell Culture: A proliferating mass of cells which may be in an undifferentiated or differentiated state.

Chimeric Sequence or Gene: A nucleotide sequence derived from at least two heterologous parts. The sequence may comprise DNA or RNA.

Coding Sequence: A deoxyribonucleotide sequence which when transcribed and translated results in the formation of a cellular polypeptide, or a ribonucleotide sequence which when translated results in the formation of a cellular polypeptide.

Compatible: The capability of operating with other components of a system. A vector which is compatible with a host is one which is capable of replicating in that host. A coat protein which is compatible with a viral nucleotide sequence is one which is capable of encapsidating the viral sequence.

Fusion Protein: An amino acid sequence which contains the parts or whole amino acid sequence of two or more heterologous proteins. An example of one means to obtain a fusion protein is the genetic modification of a host gene by inserting or attaching a nucleic acid sequence which contains a foreign gene and subsequent translation to produce the gene product. Generally the regulation sites on the host structural gene remain

functional. Thus, protein biosynthesis can occur in the usual way from the host start codon to the stop codon on the foreign gene. The fusion protein may also contain one or more amino acid residues between the heterologous peptides for separation of the peptides.

Gene: A discrete chromosomal region which is responsible for a discrete cellular product.

Host: A cell, tissue or organism capable of replicating a viral vector and which is capable of being infected by a virus containing the viral vector. This term is intended to include prokaryotic and eukaryotic cells, organs, tissues or organisms, such as bacteria, yeast, fungus, animal cells and plant tissue.

Infection: The ability of a virus to transfer its nucleic acid to a host wherein the viral nucleic acid is replicated, viral proteins are synthesized and new viral particles assembled. The terms transmissible and infective are used interchangeably herein. The term non-infective as used herein means non-infective by natural, biological means.

Phenotypic Trait: An observable property resulting from the expression of a gene.

Plant Cell: The structural and physiological unit of plants, consisting of a protoplast and the cell wall.

Plant Organ: A distinct and visibly differentiated part of a plant such as root, stem, leaf or embryo.

Plant Tissue: Any tissue of a plant in plant or in culture. This term is intended to include a whole plant, plant cell, plant organ, protoplast, cell culture or any group of plant cells organized into a structural and functional unit.

Production Cell: A cell, tissue or organism capable of replicating a vector or a viral vector, but which is not necessarily a host to the virus. This term is intended to include prokaryotic and eukaryotic cells,

organs, tissues or organisms, such as bacteria, yeast, fungus, animal cells and plant tissue.

Promoter: The 5'-flanking, non-coding sequence adjacent a coding sequence which is involved in the initiation of transcription of the coding sequence.

Protoplast: An isolated plant cell without cell walls, having the potency for regeneration into cell culture or a whole plant.

Specific Cleavage: Cleavage of a fusion protein at or near the point at which the heterologous peptides, e.g. host peptide and the foreign protein, are fused. Specific cleavage can take place at a genetically engineered linker sequence by enzyme hydrolysis, at the carboxy end of a methionine residue by reaction with cyanogan bromide, or at a naturally occuring protease sequence.

Substantial Sequence Homology: Denotes nucleotide sequences that are substantially functionally equivalent to one another. Nucleotide differences between such sequences having substantial sequence homology will be de minimus in affecting the function of the gene products or an RNA coded for by such sequence.

Transcription: The production of an RNA molecule by RNA polymerase as a complementary copy of a DNA sequence.

Vector: A self-replicating DNA molecule which transfers a DNA segment between cells.

Viral Vector: A vector comprising a nucleic acid sequence of a virus which has been modified so that a non-biologically functional coat protein is produced. This may be accomplished by removing at least a part of the coding sequence or by mutating the coding sequence. If the virus codes for one more virus transmissibility factors, then the nucleic acid sequence of the virus is also modified to make these non-biologically functional.

Virus: An infectious agent which is composed of a nucleic acid encapsidated in a protein. A virus may be a mono-, di-, tri- or multi-partite virus as described above.

The present invention provides for the infection of a host, such as a prokaryotic or eukaryotic organism, cell or tissue, by a virus which has been modified so that the virus is transmissible but the viral nucleic acid is not infective. Naturally occurring mutant viruses may also have these same properties of being transmissible, but having viral nucleic acid which is not infective. The non-infectivity of the viral nucleic acid is accomplished by modifying the nucleic acid so that biologically non-functional viral coat protein (capsid protein) and any other viral transmissibility factors are produced as described herein.

The non-infectivity of the viral nucleic acid is accomplished by deleting at least a part of the coding sequence for the viral coat protein and any other viral transmissability factors which may be present in the viral nucleic acid.

The present invention has a number of advantages, one of which is that the transformation and regeneration of target organisms is not necessary. Another advantage is that it is not necessary to develop vectors which integrate a desired coding sequence in the genome of the target organism. Existing organisms can be altered with a new coding sequence without the need of going through a germ cell. The present invention also gives the option of applying the coding sequence to the desired organism, tissue, organ or cell.

The chimeric genes and vectors of the present invention are constructed using techniques well known in the art. Suitable techniques have been described in Maniatis, T. et al., *Molecular Cloning*, Cold Spring Harbor Laboratory, New York (1982); *Methods in*

Enzymology, Vols. 68, 100, 101, 118 and 152-155, Academic Press, New York (1979, 1983, 1983, 1986 and 1987); and *DNA Cloning*, Vols. I, II, III, Glover, D.M., Ed., IRL Press, Oxford (1985 and 1987). Medium compositions have been described in Miller, J.H., *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, New York (1972), as well as the references previously identified. DNA manipulations and enzyme treatments are carried out in accordance with the manufacturers' recommended procedures.

An important feature of the present invention is the preparation of nucleotide sequences which are capable of replication in a compatible host but which in themselves are incapable of infecting the host. The nucleotide sequence has substantial sequence homology to a viral nucleotide sequence. The viral nucleotide sequence may be a prokaryotic or eukaryotic viral nucleotide sequence. Suitable viral nucleotide sequences include those of viruses which infect bacteria, yeast, fungus, animals and plants. A partial listing of suitable viruses has been described above. The nucleotide sequence may be an RNA, DNA, cDNA or chemically synthesized RNA or DNA.

The first step in achieving any of the features of the invention is to modify the nucleotide sequences coding for the capsid protein and any transmissibility factors within the viral nucleotide sequence by known conventional techniques such that non-biologically functional proteins are produced by the modified virus. Therefore, any virus for which the capsid protein nucleotide sequence and any transmissibility factor nucleotide sequences have been identified may be suitable for use in the present invention. Other viruses may be used after the nucleic acid has been sequenced.

Some of the viruses which meet this requirement, and therefore are suitable, include the alphaviruses such as Eastern Equine Encephalomyelitis virus (EEEV), Western Equine Encephalomyelitis virus (WEEV), Venezuelan Encephalomyelitis virus (VEV), Sindbis virus, Semliki Forest virus (SFV) and Ross River virus (RRV), the rhinoviruses such as human rhinovirus 2 (HRV2) and human rhinovirus type 89 (HRV89), the polioviruses such as poliovirus 2 (PV2) and poliovirus 3 (PV3), simian virus 40 (SV40), viruses from the tobacco mosaic virus group such as Tobacco Mosaic virus (TMV), Cowpea Mosaic virus (CMV), Alfalfa Mosaic virus (AmV), Cucumber Green Mottle Mosaic virus watermelon strain (CGMMV-W) and Oat Mosaic virus (OMV) and viruses from the brome mosaic virus group such as Brome Mosaic virus (BMV), broad bean mottle virus and cowpea chlorotic mottle virus. Additional suitable viruses include Rice Necrosis virus (RNV), adenovirus type 2 and geminiviruses such as tomato golden mosaic virus (TGMV), cassava latent virus and maize streak virus. Each of these groups of suitable viruses is characterized below.

ALPHAVIRUSES

The alphaviruses are a genus of viruses of the family Togaviridae. Almost all of the members of this genus are transmitted by mosquitoes, and may cause diseases in man or animals. Some of the alphaviruses are grouped into three serologically defined complexes. The complex-specific antigen is associated with the E1 protein of the virus, and the species-specific antigen is associated with the E2 protein of the virus.

The Semliki Forest virus complex includes Bebaru virus, Chikungunya Fever virus, Getah virus, Mayaro Fever virus, O'nyongnyong Fever virus, Ross River virus, Sagiyama virus, Semliki Forest virus and Una virus. The

Venezuelan Equin Encephalomyelitis virus complex includes Cabassou virus, Ev rglad s virus, Mucambo virus, Pixuna virus and Venezuelan Equine Encephalomyelitis virus. The Western Equine Encephalomyelitis virus complex includes Aura virus, Fort Morgan virus, Highlands J virus, Kyzylagach virus, Sindbis virus, Western Equine Encephalomyelitis virus and Whataroa virus.

The alphaviruses contain an icoschedral nucleocapsid consisting of 180 copies of a single species of capsid protein complexed with a plus-stranded 425 to 498 mRNA of up to about 11,703 nucleotides. The alphaviruses mature when preassembled nucleocapsid is surrounded by a lipid envelope containing two virus encoded integral membrane glycoproteins, called E1 and E2. The envelope is acquired when the capsid, assembled in the cytoplasm, buds through the plasma membrane. The envelope consists of a lipid bilayer derived from the host cell.

The 425 to 498 mRNA encodes a glycoprotein which is contranslationally cleaved into nonstructural proteins and structural proteins. The 3' one-third of the RNA genome consists of a 26S mRNA which encodes for the capsid protein and the E3, E2, K6 and E1 glycoproteins. The capsid protein is cotranslationally cleaved from the E3 protein. It is hypothesized that the amino acid triad of His, Asp and Ser at the COOH terminus of the capsid protein comprises a serine protease responsible for cleavage. Hahn, C.S. et al., *Proc. Natl. Acad. Sci. U.S.A. 82*, 4648 (1985). Cotranslational cleavage also occurs between E2 and K6 proteins. Thus two proteins PE2 which consists of E3 and E2 prior to cleavage and an E1 protein comprising K6 and E1 are formed. These proteins are cotranslationally inserted into the endoplasmic reticulum of the host cell, glycosylated and transported via the Golgi apparatus to the plasma

membrane where they can be used for budding. At the point of virion maturation the E3 and E2 proteins are separated. The E1 and E2 proteins are incorporated into the lipid envelope.

It has been suggested that the basic amino-terminal half of the capsid protein stabilizes the interaction of capsid with genomic RNA, Garoff, H. et al., *supra*; or interacts with genomic RNA to initiate encapsidation, Strauss E.G. et al., in the *Togaviruses and Flaviviruses*, Ed. S. Schlesinger & M. Schlesinger, Plenum Press, New York, p. 35-90, (1980). These suggestions imply that the origin of assembly is located either on the unencapsidated genomic RNA or at the amino-terminus of the capsid protein. It has been suggested that E3 and K6 function as signal sequences for the insertion of PE2 and E1, respectively, into he endoplasmic reticulum. Garoff, H. et al., *supra*; Delgarno, L. et al., *Virology 120*, 170 (1983).

Work with temperature sensitive mutants of alphaviruses has shown that failure of cleavage of the structural proteins results in failure to form mature virions. Lindquist, B.H. et al., *Virology 151*, 10 (1986) characterized a temperature sensitive mutant of Sindbis virus, t_s20. Temperature sensitivity results from an A - U change at nucleotide 9502. The t_s lesion present cleavage of PE2 to E2 and E3 and the final maturation of progeny virions at the nonpermissive temperature. Hahn, C.S. et al., *supra*, reported three temperature sensitive mutations in the capsid protein which prevents cleavage of the precursor polyprotein at the nonpermissive temperature. The failure of cleavage resulted in no capsid formation and very little envelope protein.

Defective interfering RNA's (DI particles) of Sindbis virus are helper-dependent deletion mutants which interfere specifically with the replication of the

homologous standard virus. Perrault, J., *Microbiol. Immunol. 93*, 151 (1981). DI particles have been found to be functional vectors for introducing at least one foreign gene into cells. Levis, R., *Proc. Natl. Acad. Sci. U.S.A. 84*, 4811 (1987).

It has been found that it is possible to replace at least 1689 internal nucleotides of a DI genome with a foreign sequence and obtain RNA that will replicate and be encapsidated. Deletions of the DI genome do not destroy biological activity. The disadvantages of the system are that DI particles undergo apparently random rearrangements of the internal RNA sequence and size alterations. Monroe, S.S. et al., *J. Virology 49*, 865 (1984). Expression of a gene inserted into the internal sequence is not as high as expected. Levis, R. et al., supra, found that replication of the inserted gene was excellent but translation was low. This could be the result of competition with whole virus particles for translation sites and/or also from disruption of the gene due to rearrangement through several passages.

Two species of mRNA are present in alphavirus-infected cells: A 42S mRNA region, which is packaged into nature virions and functions as the message for the nonstructural proteins, and a 26S mRNA, which encodes the structural polypeptides. The 26S mRNA is homologous to the 3' third of the 42S mRNA. It is translated into a 130K polyprotein that is cotranslationally cleaved and processed into the capsid protein and two glycosylated membrane proteins, E1 and E2.

The 26S mRNA of Eastern Equine Encephalomyelitis (EEE) strain 82V-2137 was cloned and analyzed by Chang et al. *J. Gen. Virol. 68*, 2129 (1987). The 26S mRNA region encodes the capsid proteins, E3, E2, 6K and E1. The amino terminal end of the capsid protein is thought to either stabilize the interaction of capsid with mRNA

or to interact with genomic RNA to initiate encapsidation.

Uncleaved E3 and E2 proteins called PE2 is inserted into the host endoplasmic reticulum during protein synthesis. The PE2 is thought to have a region common to at least five alphaviruses which interacts with the viral nucleocapsid during morphogenesis.

The 6K protein is thought to function as a signal sequence involved in translocation of the E1 protein through the membrane. The E1 protein is thought to mediate virus fusion and anchoring of the E1 protein to the virus envelope.

RHINOVIRUSES

The rhinoviruses are a genus of viruses of the family Picornaviridae. The rhinoviruses are acid-labile, and are therefore rapidly inactivated at pHs less than about 6. The rhinoviruses commonly infect the upper respiratory tract of mammals.

Human rhinoviruses are the major causal agents of the common cold, and many serotypes are known. Rhinoviruses may be propagated in various human cell cultures, and have an optimum growth temperature of about 33°C. Most strains of rhinoviruses are stable at or below room temperature and can withstand freezing. Rhinoviruses can be inactivated by citric acid, tincture of iodine or phenol/alcohol mixtures.

The complete nucleotide sequence of human rhinovirus 2 (HRV2) has been sequenced. The genome consists of 7102 nucleotides with a long open reading frame of 6450 nucleotides which is initiated 611 nucleotides from the 5' end and stops 42 nucleotides from the poly(A) tract. Three capsid proteins and their cleavage cites have been identified.

Rhinovirus RNA is singl -stranded and positiv sense. The RNA is not capped, but is joined at the 5' end to a small virus encoded protein , virion-protein genome-linked (VPg). Translation is presumed to result in a single polyprotein which is broken by proteolytic cleavage to yield individual virus proteins.

An icosahedral viral capsid contains 60 copies each of 4 virus proteins VP1, VP2, VP3 and VP4 and surrounds the RNA genome. Medappa, K. C. et al. *Virology 44*, 259 (1971).

Analysis of the 610 nucleotides preceding the long open reading frame shows several short open reading frames. However, no function can be assigned to the translated proteins since only two sequences show homology throughout HRV2, HRV14 and the 3 serotypes of poliovirus. These two sequences may be critical in the life cycle of the virus. They are a stretch of 16 bases beginning at 436 in HRV2 and a stretch of 23 bases beginning at 531 in HRV2. Cutting or removing these sequences from the remainder of the sequence for non-structural proteins could have an unpredictable effect upon efforts to assemble a mature virion.

The capsid proteins of HRV2: VP4, VP2, VP3 and VP1 begin at nucleotide 611, 818, 1601 and 2311, respectively. The cleavage point between VP1 and P2A is thought to be around nucleotide 3255. Skern, T. et al., *Nucleic Acids Research 13*, 2111 (1985).

Human rhinovirus type 89 (HRV89) is very similar to HRV2. It contains a genome of 7152 nucleotides with a single large open reading frame of 2164 codons. Translation begins at nucleotide 619 and ends 42 nucleotides before the poly(A) tract. The capsid structural proteins, VP4, VP2, VP3 and VP1 are the first to be translated. Translation of VP4 begins at 619. Cleavage cites occur at:

VP4/VP2	825	determined
VP2/VP3	1627	determined
VP3/VP1	2340	determined
VP1/P2-A	3235	presumptive

Duechler, M. et al., *Proc. Natl. Acad. Sci. USA 84*, 2605 (1987).

POLIOVIRUSES

Polioviruses are the causal agents of poliomyelitis in man, and are one of three groups of enteroviruses. Enteroviruses are a genus of the family Picornaviridae (also the family of rhinoviruses). Most enteroviruses replicate primarily in the mammalian gastrointestinal tract, although other tissues may subsequently become infected. Many enteroviruses can be propagated in primary cultures of human or monkey kidney cells and in some cell lines (e.g. HeLa, Vero, WI-38). Inactivation of the enteroviruses may be accomplished with heat (about 50°C), formaldehyde (3%), hydrochloric acid (0.1N) or chlorine (ca. 0.3-0.5 ppm free residual Cl_2).

The complete nucleotide sequence of poliovirus PV2 (Sab) and PV3 (Sab) have been determined. They are 7439 and 7434 nucleotide in length, respectively. There is a single long open reading frame which begins more than 700 nucleotides from the 5' end. Poliovirus translation produces a single polyprotein which is cleaved by proteolytic processing. Kitamura, N. et al. *Nature, 291*, 547 (1981).

It is speculated that these homologous sequences in the untranslated regions play an essential role in viral replication such as:

1. viral-specific RNA synthesis;
2. viral-specific protein synthesis; and
3. packaging

Toyoda, H. et al., *J. Mol. Biol., 174,* 561 (1984).

The structures of the serotypes of poliovirus have a high degree of sequence homology. Their coding sequences the same proteins in the same order. Therefore, genes for structural proteins are similarly located. In PV1, PV2 and PV3, the polyprotein begins translation near the 750 nucleotide. The four structural proteins VP4, VP2, VP3 and VP1 begin at about 745, 960, 1790 and 2495, respectively, with VPI ending at about 3410. They are separated *in vivo* by proteolytic cleavage, rather than by stop/start codons.

SIMIAN VIRUS 40

Simian virus 40 (SV40) is a virus of the genus Polyomavirus, and was originally isolated from the kidney cells of the rhesus monkey. The virus is commonly found, in its latent form, in such cells. Simian virus 40 is usually non-pathogenic in its natural host.

Simian virus 40 virions are made by the assembly of three structural proteins, VP1, VP2 and VP3. Girard, M. et al., *Biochem. Biophys. Res. Commun. 40,* 97 (1970); Prives, C. L. et al., *Proc. Natl. Acad. Sci. USA 71,* 302 (1974); and Rozenblatt, S. et al., *Proc. Natl. Acad. Sci. USA 73,* 2747 (1976). The three corresponding viral genes are organized in a partially overlapping manner. They constitute the late genes portion of the genome. Tooze, J., *Molecular Biology of Tumor Viruses,* 2nd Ed. Part 2, p. 799-831 (1980). Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. Capsid proteins VP2 and VP3 are encoded by nucleotides 545 to 1601 and 899 to 1601, respectively, and both are read in the same frame. VP3 is therefore a subset of VP2. Capsid protein VP1 is encoded by nucleotides 1488-2574. The end of the VP2-VP3 open reading frame therefore overlaps

the VP1 by 113 nucleotides but is read in an alternative frame. Tooze, J., *supra*. Sychowski, C. et al., *J. Virology 61*, 3862 (1987)

TOBACCO MOSAIC VIRUS GROUP

Tobacco Mosaic virus (TMV) is a type member of the Tobamoviruses. The TMV virion is a tubular filament, and comprises coat protein subunits arranged in a single right-handed helix with the single-stranded RNA intercalated between the turns of the helix. TMV infects tobacco as well as other plants. TMV is transmitted mechanically and may remain infective for a year or more in soil or dried leaf tissue.

The TMV virions may be inactivated by subjection to an environment with a pH less than 3 or greater than 8, or by formaldehyde or iodine. Preparations of TMV may be obtained from plant tissues by $(NH_4)_2SO_4$ precipitation followed by differential centrifugation.

The TMV single-stranded RNA genome is about 6400 nucleotides long and is capped at the 5' end but is not poly-adenylated. The genomic RNA can serve as mRNA for a protein of molecular weight about 130,000 (130K) and another produced by read-through of molecular weight about 180,000 (180K). However, it cannot function as a messenger for the synthesis of coat protein. Other genes are expressed during infection by the formation of monocistronic, 3'-coterminal subgenomic mRNAs, including one (LMC) encoding the 17.5K coat protein and another (I_2) encoding a 30K protein. The 30K protein has been detected in infected protoplasts (*Virology 132*, 71 (1984)), and it is involved in the cell-to-cell transport of the virus in an infected plant (Deom, C.M. et al., *Science 237*, 389 (1987)). The functions of the two large proteins are unknown.

Several double-stranded RNA mol cules, including double-stranded RNAs corresponding to the genomic, I_2 and LMC RNAs, have been detected in plant tissues infected with TMV. These RNA molecules are presumably intermediates in genome replication and/or mRNA synthesis - processes which appear to occur by different mechanisms.

TMV assembly apparently occurs in the plant cell cytoplasm, although it has been suggested that some TMV assembly may occur in chloroplasts since transcripts of ctDNA have been detected in purified TMV virions. Initiation of TMV assembly occurs by interaction between ring-shaped aggregates ("discs") of coat protein (each disc consisting of two layers of 17 sub-units) and a unique internal nucleation site in the RNA; a hairpin region about 900 nucleotides from the 3' end in the common strain of TMV. Any RNA, including subgenomic RNAs, containing this site may be packaged into virions. The discs apparently assume a helical form on interaction with the RNA, and assembly (elongation) then proceeds in both directions (but much more rapidly in the 3'- to 5' direction from the nucleation site).

Another member of the Tobamoviruses, the Cucumber green mottle mosaic virus watermelon strain (CGMMV-W) is related to the cucumber virus. Noru, Y., et al., Virology 45, 577 (1971). The coat protein of CGMMV-W interacts with the RNA of both TMV and CGMMV to assemble viral particles *in vitro*. Kurisu et al., *Virology 70*, 214 (1976).

Several strains of the tobamovirus group are divided into two subgroups on the basis of the location of the assembly of origin. Fukuda, M. et al., *Proc. Natl. Acad. Sci. USA 78*, 4231 (1981). Subgroup I, which includes the vulgare, OM, and tomato strain, has an origin of assembly at about 800-1000 nucleotides from the 3' end of the RNA genome, and outside of the coat

protein cistron. Lebeurier, G. et al., *Proc. Natl. Acad. Sci. USA 74*, 1913 (1977); and Fukuda, M. et al., Virology 101, 493 (1980). Subgroup II, which includes CGMMV-W and cornpea strain (Cc), has an origin of assembly about 300-500 nucleotides from the 3' end of the RNA genome, and within the coat-protein cistron. Fukuda, M. et. al., *supra*. The coat protein cistron of CGMMV-W is located at nucleotides 176-661 from the 3'end. The 3' noncoding region is 175 nucleotides long. The origin of assembly is positioned within the coat protein cistron. Meshi, T. et al., *Virology 127*, 52 (1983).

BROME MOSAIC VIRUS GROUP

Brome mosaic virus (BMV) is a member of a group of tripartite single-stranded RNA-containing plant viruses commonly referred to as the bromoviruses. Each member of the bromoviruses infects a narrow range of plants. Mechanical transmission of bromoviruses occurs readily, and some members are transmitted by beetles. In addition to BMV, other bromoviruses include broad bean mottle virus and cowpea chlorotic mottle virus.

Typically, a bromovirus virion is icosahedral with a diameter of about 26 mm., and contains a single species of coat protein. The bromovirus genome has three molecules of linear, positive-sense, single-stranded RNA, and the coat protein mRNA is also encapsidated. The RNAs each have a capped 5' end and a tRNA-like structure (which accepts tyrosine) at the 3' end. Virus assembly occurs in the cytoplasm. The complete nucleotide sequence of BMV has been identified and characterized as described by Alquist et al., *J. Mol. Biol. 153*, 23 (1981).

RICE NECROSIS VIRUS

Rice Necrosis virus is a member of the Potato Virus Y Group or Potyviruses. The Rice Necrosis virion is a flexuous filament comprising one type of coat protein (molecular weight about 32,000 to about 36,000) and one molecule of linear positive-sense single-standard RNA. The Rice Necrosis virus is transmitted by *Polymyxa graminis* (a eukaryotic intracellular parasite found in plants, algae and fungi).

ADENOVIRUSES

Adenovirus type 2 is a member of the adenovirus family or adenovirus. This family of viruses are non-enveloped, icosahedral, linear, double-stranded DNA-containing viruses which infect mammals or birds.

The adenovirus virion consists of an icosahedral capsid enclosing a core in which the DNA genome is closely associated with a basic (arginine-rich) viral polypeptide VII. The capsid is composed of 252 capsomeres: 240 hexons (capsomers each surrounded by 6 other capsomers) and 12 pentons (one at each vertex, each surrounded by 5 'peripentonal' hexons). Each penton consists of a penton base (composed of viral polypeptide III) associated with one (in mammalian adenoviruses) or two (in most avian adenoviruses) glycoprotein fibres (viral polypeptide IV). The fibres can act as haemagglutinins and are the sites of attachment of the virion to a host cell-surface receptor. The hexons each consist of three molecules of viral polypeptide II; they make up the bulk of the icosahedron. Various other minor viral polypeptides occur in the virion.

The adenovirus dsDNA genome is covalently linked at the 5' end of each strand to a hydrophobic 'terminal

protein', TP (molecular weight about 55,000); the DNA has an inverted terminal repeat of different length in different adenoviruses. In most adenoviruses examined, the 5'-terminal residue is dCMP.

During its replication cycle, the virion attaches via its fibres to a specific cell-surface receptor, and enters the cell by endocytosis or by direct penetration of the plasma membrane. Most of the capsid proteins are removed in the cytoplasm. The virion core enters the nucleus, where the uncoating is completed to release viral DNA almost free of virion polypeptides. Virus gene expression then begins. The viral dsDNA contains genetic information on both strands. Early genes (regions E1a, E1b, E2a, E3, E4) are expressed before the onset of viral DNA replication. Late genes (regions L1, L2, L3, L4 and L5) are expressed only after the initiation of DNA synthesis. Intermediate genes (regions E2b and IVa_2) are expressed in the presence or absence of DNA synthesis. Region E1a encodes proteins involved in the regulation of expression of other early genes, and is also involved in transformation. The RNA transcripts are capped (with $m^7G^5ppp^5N$) and polyadenylated in the nucleus before being transferred to the cytoplasm for translation.

Viral DNA replication requires the terminal protein, TP, as well as virus-encoded DNA polymerase and other viral and host proteins. TP is synthesized as an 80K precursor, pTP, which binds covalently to nascent replicating DNA strands. pTP is cleaved to the mature 55K TP late in virion assembly; possibly at this stage, pTP reacts with a dCTP molecule and becomes covalently bound to a dCMP residue, the 3' OH of which is believed to act as a primer for the initiation of DNA synthesis. Late gene expression, resulting in the synthesis of viral structural proteins, is accompanied by the cessation of cellular protein synthesis, and virus

assembly may result in the production of up to 10^5 virions per cell.

GEMINIVIRUSES

Geminiviruses are a group of small, single-stranded DNA-containing plant viruses with virions of unique morphology. Each virion consists of a pair of isometric particles (incomplete icosahedra), composed of a single type of protein (molecular weight about $2.7-3.4 \times 10^4$). Each geminivirus virion contains one molecule of circular, positive-sense, single-stranded DNA. In some geminiviruses (i.e., cassava latent virus and bean golden mosaic virus), the genome appears to be bipartite, containing two single-stranded DNA molecules which are of similar size, but differ as to nucleotide sequence. However, other geminiviruses (i.e., the leafhopper transmitted viruses such as *Chloris* striate mosaic virus), have only one type of single-stranded DNA. Geminivirus replication occurs in the plant cell nucleus where large aggregates of virus particles accumulate.

The nucleotide sequence of any suitable virus can be derived from a viral nucleic acid by modifying the coat protein coding sequence. The modification may be the removal of a coding sequence for at least a part of the viral coat protein. Alternatively, the nucleotide sequence can be synthesized such that it lacks at least a part of the viral coat protein coding sequence. A sufficient amount of the coding sequence is removed such that any coat protein which may be produced by the virus will be incapable of encapsidating the viral nucleic acid. In addition, the coat protein coding sequence may be modified by mutation such that the coat protein which is produced is incapable of encapsidating the viral

nucleic acid. In each instance, as non-biologically functional protein is produced. In some instanc s, part of the gene may be necessary for good promoter activity. If so, then the entire gene should not be deleted. In order to be easily transmissible to other plants, the nucleotide sequence must be encapsidated in a compatible coat protein as described further below. The viral nucleic acid may further be modified to alter the coding sequence for any viral transmissibility factors. The alteration of the coding sequences for these factors will ensure that the nucleotide sequence cannot be transmitted by other vectors, e.g. by insects.

The nucleotide sequence is prepared by cloning the viral nucleic acid in an appropriate production cell. If the viral nucleic acid is DNA, it can be cloned directly into a suitable vector using conventional techniques. One technique is to attach an origin of replication compatible with the production cell to the viral DNA. If the viral nucleic acid is RNA, a full-length DNA copy of the viral genome is first prepared by well known procedures. For example, the viral RNA is transcribed into DNA using reverse transcriptase to produce subgenomic DNA pieces, and a double-stranded DNA made using DNA polymerases. The DNA is then cloned into appropriate vectors and cloned into a production cell. The DNA pieces are mapped and combined in proper sequence to produce a full-length DNA copy of the viral RNA genome. The coding sequences for the viral coat protein and any viral transmissibility factors are identified and altered. The resulting nucleotide sequence is self-replicating but incapable of infecting host itself. Any manner of separating the coding sequences for the viral coat protein and the coding sequences for the viral transmissibility factors from the remainder of the viral nucleotide sequence or of altering the nucleotide sequences of these proteins to

produce non-biologically functional prot ins is suitabl for the present invention.

In the case of alphaviruses, the E1 and E2 glycoproteins may play a role in transmissibility of the virus (Garaff, H. et al., *Nature 228*, 236 (1980)). These glycoproteins are incorporated in a lipid envelope which surrounds the coat protein. The nucleotide sequence which codes for the E1 and E2 glycoproteins is adjacent to the coding sequence for the coat protein in alphavirus RNA. Therefore the E1 and E2 glycoprotein coding sequences can be removed with the coat protein coding sequence by known conventional techniques.

A second feature of the present invention is a chimeric nucleotide sequence which comprises a first nucleotide sequence and a second nucleotide sequence. The first sequence is capable of self-replication, is not capable of transmission and has substantial sequence homology to a viral nucleotide sequence, as described above. The second sequence is capable of being transcribed in a host. The second sequence is preferably placed adjacent a viral promoter, although a fusion protein may be produced which also has biological activity. Any viral promoter can be utilized, but it is preferred to use a promoter of the viral coat protein gene, at least a part of the coding sequence of which has been deleted. In those instances where the coat protein coding sequence is altered but not deleted, a viral promoter can be attached to the second sequence by conventional techniques or the second sequence can be inserted into or adjacent the coat protein coding sequence such that a fusion protein is produced. The second sequence in the chimeric nucleotide sequence is transcribed and translated in the host, e.g., a microorganism or animal or plant tissue, to produce a stereospecific enzyme. The chimeric nucleotide sequence is constructed using conventional techniques such that

the second nucleotide sequence is in proper orientation and reading frame to the viral promot r.

A fusion protein can be formed by incorporation of the second sequence into a structural gene of the viral nucleic acid, e.g., the coat protein gene. The regulation sites on the viral structural gene remain functional. Thus, protein synthesis can occur in the usual way from the starting codon for methionine to the stop codon on the foreign gene to produce the fusion protein. The fusion protein contains at the amino terminal end, a part of all of the viral structural protein at the amino terminal end and at the carboxy terminal end, the desired stereospecific enzyme. For its subsequent use, the stereospecific enzyme must first be processed by a specific cleavage from this fusion protein. A reaction with cyanogen bromide leads to a cleavage of the peptide sequence at the carboxy end of methionine residues (S.B. Needleman, "Protein Sequence Determination", Springer Publishers, 1970, N.Y.). Accordingly, it is necessary for this purpose that the second sequence contain an additional codon for methionine, whereby a methionine residue is disposed between the N-terminal native protein sequence and the stereospecific enzyme of the fusion protein. This method, however, fails if other methionine residues are present in the desired stereospecific enzyme. Additionally, the cleavage with cyanogen bromide has the disadvantage of evoking secondary reactions at various other amino acids.

Alternatively an oligonucleotide segment, referred to as a "linker" may be placed between the second sequence and the viral sequence. The linker codes for an amino acid sequence of the extended specific cleavage site of a proteolytic enzyme as well as a specific cleavage site. For example, U.S. Patent Nos. 4,769,326 and 4,543,329. The use of linkers in the fusion protein

at the amino terminal end of the stereospecific nzyme sequence avoids the secondary reactions inher nt in cyanogen bromide cleavage, by a selective enzymatic hydrolysis. An example of such a linker is a tetrapeptide of the general formula Pro-Xyz-Gly-Pro-(amino-terminal end of stereospecific enzyme), wherein Xyz is any desired amino acid. The overall cleavage is effected by first selectively cleaving the Xyz-Gly bond with a collagenase (E.C. 3.4.24.3., Clostridiopeptidase A) then removing the glycine residue with an aminoacylproline aminopeptidase (aminopeptidase-P, E.C. 3.4.11.9.) and removing the proline residue with a proline aminopeptidase (E.C. 3.4.11.5). In the alternative, the aminopeptidase enzyme can be replaced by postproline dipeptidylaminopeptidase. Other linkers and appropriate enzymes are set forth in U.S. Patent No. 4,769,326.

The fusion protein is produced and initially purified. The stereospecific enzyme is then removed by the appropriate cleavage and further purified.

Fusion proteins are produced preferably to protect the foreign gene product from degradation within the host. Where it is desirable to produce a gene product that is active within the host a coding sequence for an amino acid sequence such as the triad His, Asp and Ser, which occurs at the carboxyl terminus of the capsid protein of an alphavirus, may be used. It is believed that this particular sequence codes for a serine protease which results in cotranslational cleavage. Hahn, C.S. et al. *Supra*.

An example of a chimeric nucleotide sequence is one which contains a first nucleotide sequence having substantial sequence homology to TMV and a second nucleotide sequence which is a coding sequence for an esterase for the stereospecific hydrolysis of α-methylarylacetic acid esters. In a second example, the

virus is oat mosaic virus (OMV) or rice necrosis virus (RNV). OMV and RNV are capable of infecting most monocot species including, but not limited to, barley and corn. In a third example, the second nucleotide sequence is the coding sequence for a lipase for the stereospecific hydrolysis of α-methylarylacetic acid esters. Potato virus Y (PVY) or potato virus X (PVX) is used in a fourth example.

The second nucleotide sequence can be inserted into the first nucleotide sequence prepared above such that it is adjacent a viral promoter. Since the location of the promoter of the viral coat protein gene is known in this sequence as a result of the deletion of the gene, the second nucleotide sequence can be placed adjacent this promoter. Alternatively, an appropriate viral promoter can first be attached to the second nucleotide sequence and this construct can then be inserted either into the first nucleotide sequence or adjacent thereto. In addition, the second nucleotide sequence can be inserted into and adjacent an altered coat protein coding sequence.

A double-stranded DNA of the chimeric nucleotide sequence or of a complementary copy of the chimeric nucleotide sequence is cloned into a production cell. If the viral nucleic acid is an RNA molecule, the chimeric nucleotide sequence is first attached to a promoter which is compatible with the production cell. The chimeric nucleotide sequence can then be cloned into any suitable vector which is compatible with the production cell. In this manner, only RNA copies of the chimeric nucleotide sequence are produced in the production cell. For example, if the production cell is *E. coli*, the *lac* promoter can be utilized. If the production cell is a plant cell, the CaMV promoter can be used. The production cell will be a eukaryotic cell such as yeast, plant or animal, if viral RNA must be

capped for biological activity. The chimeric nucleotide sequence can th n be cloned into any suitable vector which is compatible with the production cell. Alternatively, the chimeric nucleotide sequence is inserted in a vector adjacent a promoter which is compatible with the production cell. If the viral nucleic acid is a DNA molecule, it can be cloned directly into a production cell by attaching it to an origin of replication which is compatible with the production cell. In this manner, DNA copies of the chimeric nucleotide sequence are produced in the production cell.

In the case of alphaviruses, where the E1 and E2 glycoproteins and the coat protein are removed, a larger foreign protein coding sequence may be inserted to form the chimeric nucleotide sequence. The E1 and E2 glycoproteins do not have their own promoters, so the coat protein promoter would be used for a foreign coding sequence inserted in place of the adjacent coding sequences for the coat protein, E1 glycoprotein and E2 glycoprotein.

Alternatively, more than one foreign coding sequence may be inserted in place of the adjacent coding sequences for the coat protein, E1 glycoprotein and E1 glycoprotein. However, in this case, each foreign coding sequence would require its own appropriate viral promoter. However, the coat protein promoter could be used for one foreign coding sequence if that promoter had been preserved in the nucleotide sequence.

A promoter is a DNA sequence that directs RNA polymerase to bind to DNA and to initiate RNA synthesis. There are strong promoters and weak promoters. Among the strong promoters are lacuv5, trp, tac, trp-lacuv5, λp_1, ompF, and bla. A useful promoter for expressing foreign genes in *E. coli* is both strong and regulated. The λp_1 promoter of bacteriophage λ is a strong, well-

regulated promoter. Hedgpeth, J.M. et al. *Mol. Gen. Genet. 163*, 197 (1978); Bernard, H. M. et al. *Gene 5*, 59 (1979); Remaut, E.P. et al., *Gene 15*, 81 (1981).

A gene encoding a temperature-sensitive λ repressor such as λ*cIts* 857 may be included in the cloning vector. Bernard et al., *supra*. At low temperature (31°C), the p_l promoter is maintained in a repressed state by the cI-gene product. Raising the temperature destroys the activity of the repressor. The p_l promoter then directs the synthesis of large quantities of mRNA. In this way, *E. coli* production cells may grow to the desired concentration before producing the products encoded within the vectors. Similarly, a temperature-sensitive promoter may be activated at the desired time by adjusting the temperature of the culture.

It may be advantageous to assemble a plasmid that can conditionally attain very high copy numbers. For example, the pAS2 plasmid containing a *lac* or *tac* promoter will achieve very high copy numbers at 42°C. The *lac* repressor, present in the pAS2 plasmid is then inactivated by isopropyl-β-D-thiogalactoside to allow synthesis of mRNA.

A further alternative when creating the chimeric nucleotide sequence, is to prepare more than one nucleotide sequence (i.e., prepare the nucleotide sequences necessary for a multipartite viral vector construct). In this case, each nucleotide sequence would require its own origin of assembly. Each nucleotide sequence could be chimeric (i.e., each nucleotide sequence has a second sequence inserted therein), or only one of the nucleotide sequences may be chimeric.

If a multipartite virus were found to have the coding sequence for its coat protein on one strand of nucleic acid, and the coding sequence for a transmissibility factor on a different strand, then two

chimeric nucl otide strands could be created in accordance with the invention. One second sequence could be inserted in place of the coat protein gene (or inserted next to the altered coat protein gene) on one strand of viral nucleic acid, and another second sequence (coding for the same or different stereospecific enzyme) could be inserted in place of the transmissibility factor gene (or inserted next to the altered transmissibility factor gene) on the other strand of viral nucleic acid.

Alternatively, the insertion of a second sequence into the nucleotide sequence of a monopartite virus may result in the creation of two nucleotide sequences (i.e., the nucleic acid necessary for the creation of a bipartite viral vector). This would be an advantageous situation when it is desirable to keep the replication and translation of the second sequence separate from the replication and translation of some of the coding sequences of the original nucleotide sequence. Each nucleotide sequence would have to have its own origin of assembly.

A third feature of the present invention is a virus or viral particle. The virus comprises a chimeric nucleotide sequence as described above which has been encapsidated. The resulting product is then capable of infecting an appropriate host, but the chimeric nucleotide sequence is incapable of further infection since it lacks the mechanism to produce additional virus or viral particles. However, the chimeric nucleotide sequence is capable of replicating in the host. The chimeric nucleotide sequence is transcribed and/or translated within the host to produce the desired product.

Most viruses are encapsidated by either an icosahedral capsid, a spherical capsid or a rod-shaped capsid. Plant viruses such as the tobacco mosaic virus

are commonly encapsidated by a rod-shaped capsid, while alphaviruses are encapsidated by an icosahedral capsid.

The icosahedral capsids and the spherical capsids are more geometrically constrained than a rod-shaped capsid, and therefore are more limited as to the amount of nucleic acid which may be encapsidated. In contrast, a rod-shaped capsid is expandable so that it can encapsidate any amount of nucleic acid as explained in European Patent Application 0 278 667.

In one embodiment of the present invention, the chimeric nucleotide sequence is encapsidated by a heterologous capsid. Most commonly, this embodiment will make use of a rod-shaped capsid because of its ability to encapsidate a longer chimeric nucleotide sequence than the more geometrically constrained icosahedral capsid or spherical capsid. The use of a rod-shaped capsid permits the incorporation of a larger foreign protein coding sequence to form the chimeric nucleotide sequence. Such a rod-shaped capsid is most advantageous when more than one second sequence is present in the chimeric nucleotide sequence.

Another feature of the invention is a vector containing the chimeric nucleotide sequence as described above. The chimeric nucleotide sequence is adjacent a nucleotide sequence selected from the group consisting of a production cell promoter or an origin of replication compatible with the production cell. The vector is utilized to transform a production cell which will then produce the chimeric nucleotide sequence in quantity. The production cell may be any cell which is compatible with the vector. The production cell may be prokaryotic or eukaryotic. However, if the viral RNA (chimeric nucleotide sequence) must be capped in order to be active, then the production cell must be capable of capping the viral RNA, such as a eukaryotic production cell. The production cell may contain only

th vector which contains the chim ric nucleotide sequence. In this instance, the production cell will only produce the chimeric nucleotide sequence which must then be encapsidated *in vitro* to produce an infective virus. The encapsidation is accomplished by adding a compatible viral coat protein to the chimeric nucleotide sequence in accordance with conventional techniques. In this scenario, the viral coat protein would be produced in a second production cell. A second vector comprising a coding sequence for a compatible viral coat protein adjacent a promoter compatible with the production cell is utilized to transform the second production cell. The second production cell then produces the viral coat protein in quantity.

As explained above, the viral coat protein produced by the second production cell may be homologous or heterologous to the viral nucleotide sequence. The viral coat protein must be compatible with the chimeric nucleotide sequence and of sufficient size to completely encapsidate the chimeric nucleotide sequence.

Alternatively, the first production cell further contains a second vector which comprises a coding sequence for a compatible viral coat protein adjacent a promoter compatible with the production cell. Again, this coding sequence for the viral coat protein may be homologous or heterologous to the viral nucleotide sequence. In this instance, the production cell then produces the coat protein and the chimeric nucleotide sequence. The chimeric nucleotide sequence is then encapsidated by the viral coat protein to produce a transmissible virus. The second vector is prepared by conventional techniques, by inserting the coding sequence for the viral coat protein which was deleted in the preparation of the chimeric nucleotide sequence, or a homologous or heterologous viral coat protein coding

sequence into an appropriate vector under the control of a promoter compatible with the production cell.

There are numerous alternatives whereby different vectors are utilized to transform separate production cells or the same production cell. With most viruses, it is unknown whether transmissibility factors alone are essential for the infectivity of a virion or whether the transmissibility factors simply serve as an adjunct to the capsid, and therefore enhance virion infectivity, but do not prevent infection if not present on the virion. For purposes of this invention (i.e., to insure non-infectiousness of the viral vector), it has been assumed that the transmissibility factors alone could render the viral vector infectious, and therefore the vector comprising the chimeric nucleotide sequence will contain neither a nucleotide sequence for the capsid protein nor a nucleotide sequence for any transmissibility factor.

However, this vector comprising the chimeric nucleotide sequence must contain a promoter and an origin of replication, and preferably contains an origin of assembly. As stated above, this vector is utilized to transform a production cell. The origin of replication must be compatible with the production cell.

The other necessary vector or vectors may be varied primarily dependent on the virus of interest. If the viral nucleotide sequence for the capsid protein is adjacent the nucleotide sequence for the transmissibility factors, and the sequences share the same promoter (as in the case with alphaviruses), a second vector comprises the sequence for the capsid protein, the sequence for the transmissibility factors and a promoter. This second vector may optionally contain the origin of assembly if the origin of assembly has not been incorporated in the first vector. When this vector is utilized to transform a production cell,

the origin of assembly must be compatible with that pr duction cell.

Alternatively, the second vector may comprise the nucleotide sequence for the capsid protein and a promoter, and a third vector may comprise the nucleotide sequence for the transmissibility factors and a promoter. Either the second or third vector could further comprise the origin of assembly which would have to be compatible with the production cell to be transformed by the vector.

With any of the alternative embodiments, each vector may be utilized to transform its own production cell, or all of the vectors may be utilized in one production cell. All of the promoters associated with each vector must also be compatible with the production cell in which the vector is utilized.

As previously discussed, a prokaryotic production cell can be transformed to contain a vector containing the coat protein coding sequence or any other coding sequence necessary for infectivity of the virus. Alternatively, the prokaryotic production cell or a eukaryotic production cell is transformed such that the coding sequences for the coat protein and transmissibility factors (if necessary) are stably incorporated into the genome of the cell. Several conventional techniques can be utilized to accomplish this embodiment. For example, appropriate coding sequences can be inserted into plant cells by transformation with *Agrobacterium*, such as described by Schell, J. et al., *Bio/Technology 1*, 175 (1983); Fillatti, J. et al., *Bio/Technology 5*, 726 (1987); Everett, N.P. et al., *Bio/Technology 5*, 1201 (1987); Pua, E-C., *Bio/Technology 5*, 815 (1987); Hinchee, M.A., et al., *Bio/Technology 6*, 915 (1988) and *Meth. Enzymol*, Vol. 118, *supra*. *Agrobacterium* have also been utilized for introducing viruses into plants, both dicots and

monocots by a process termed Agro-infection. This technique has been described by Grimsley, N., et al., *Proc. Natl. Acad. Sci. USA 83*, 3282 (1986); Elmer, J.S. et al., *Plant Mol Biol 10*, 225 (1988); Grimsley, N. et al., *Nature 325*, 177 (1987); and Hayes, R.J. et al., *Nature 334*, 180 (1988).

Alternatively, the appropriate coding sequence can be inserted into eukaryotic cells by direct gene transfer including electroporation, calcium chloride or polyethylene glycol mediated transformation, liposome fusion microinjection or microprojectile bombardment. These techniques have been described by Fromm, M.E., *Meth. Enzymol 153*, 307 (1987); Shillito, R.D. et al., *Meth. Enzymol 153*, 283 (1987); Dehayes, A., et al., *EMBO J 4*, 2731 (1985); Negrutin, R. et al., *Plant Mol Biol 8*, 363 (1987); Reich, T.J. et al., *Bio/Technology 4*, 1001 (1986); Klein, T.M. et al., *Bio/Technology 6*, 559 (1988); McCabe, D.E. et al., *Bio/Technology 6*, 923 (1988).

A production cell which has been transformed will contain the coat protein and/or transmissibility factors coding sequence(s) either in a stable vector or stably incorporated in the genome. The vector containing the chimeric nucleotide sequence is then introduced into the production cell as previously discussed.

Once the chimeric nucleotide sequence is replicated and the capsid protein and transmissibility factors have been produced by the appropriate production cells, the intact virions may be assembled *in vitro*. The replicated chimeric nucleotide sequence is encapsidated by the capsid protein in accordance with known conventional techniques. The transmissibility factors may be added to the virion assembly separately when a third vector is used to cause the production of transmissibility factors separate from the capsid protein. Alternatively, if a single second vector was

used to cause the production of both capsid protein and transmissibility factors (as is usually the case with alphaviruses), then the *in vitro* virion assembly is complete upon encapsidation of the chimeric nucleotide sequence.

A further feature of the present invention is a host which has been infected by the virus. After introduction into a host, the host contains the chimeric nucleotide sequence which is capable of self-replication but which is not capable of producing additional transmissible viruses or viral particles. The host can be infected with the virus by conventional techniques. Suitable techniques include, but are not limited to, leaf abrasion, abrasion in solution, high velocity water spray and other injury of a host as well as imbibing host seeds with water containing the chimeric nucleotide sequence alone or the encapsidated virus. Although the chimeric nucleotide sequence is not capable of producing infective agents, it is capable of self-replicating and spreading throughout the host, e.g. if the host is a plant or animal.

An alternative method for introducing a chimeric nucleotide sequence into a plant host is a technique known as agroinfection or *Agrobacterium*-mediated transformation (sometimes called Agro-infection) as described by Grimsley, N. et al., *Nature 325*, 177 (1987). This technique makes use of a common feature of *Agrobacterium* which colonize plants by transferring a portion of their DNA (the T-DNA) into a host cell, where it becomes integrated into nuclear DNA. The T-DNA is defined by border sequences which are 25 base pairs long, and any DNA between these border sequences is transferred to the plant cells as well. The insertion of a chimeric viral nucleotide sequence between the T-DNA border sequences results in transfer of the chimeric sequence to the plant cells, where the chimeric sequence

is replicated, and th n spreads systemically through the plant. Agro-infection has been accomplished with potato spindle tuber viroid (PSTV) (Gardner, R.C. et al., *Plant Mol. Biol. 6*, 221 (1986)), cauliflower mosaic virus (CaMV) (Grimsley, N. et al., *Proc. Natl. Acad. Sci. U.S.A. 83*, 3282 (1986)), maize streak virus (Grimsley, N. et al., *Nature 325*, 177 (1987) and Lazarowitz, S.G., *Nucl. Acids Res. 16*, 229 (1988)), digitaria streak virus (Donson, J. et al., *Virogology 162*, 248 (1988)), wheat dwarf virus (Hayes, R.J. et al., *J. Gen. Virol. 69*, 891 (1988)) and tomato golden mosaic virus (TGMV) (Elmer, J.S. et al., *Plant Mol. Biol. 10*, 225 (1988) and Gardiner, W.E. et al., *EMBO J. 7*, 899 (1988)). Therefore, agro-infection of a susceptible plant could be accomplished with a virion containing a chimeric nucleotide sequence based on the nucleotide sequence of any of the above viruses.

Several viral products are useful to insure the production and spread of viral nucleic acids. One factor is a replicase which is involved in the replication of the viral nucleic acid. A second factor which may be present is termed herein a transport protein. This protein(s) is(are) involved in the movement of the viral nucleic acid from infected cells to adjacent cells and thus, the spread of the viral nucleic acid throughout the production or host cell, tissue or organism. In order to insure suitable replication and spread of the viral vector, an additional embodiment of the present invention includes the stable transformation of the host or production cell, tissue or organism with a coding sequence for viral replicase, a viral transport protein or both. The transformation, including stable integration into the genome of the production cell or host, is accomplished as described above concerning the coat protein gene.

In an additional embodiment of the invention, the host rang of a virus is enlarged so that a viral construct, i.e., chimeric nucleic acid, containing different foreign genes may be used in many different hosts. The host range of the virus is enlarged by transforming the host (or production cell since it will also work in this context) to contain parts of viral material useful in enabling infection of the host by the viral construct. For animal or bacterial hosts, the hosts are transformed to contain coding sequence(s) for the host cell receptors recognized by the virus. Alternatively, the virus can be produced to contain coding sequences of the recognition or binding proteins of the host. For plant hosts, the hosts are transformed to contain the viral replicase gene or transport protein gene. Alternatively, the virus can be produced to contain coding sequences for transport proteins capable of functioning in the host. In addition, it is possible to transform protoplasts of the host plant to contain the viral nucleic acid or parts thereof such that all cells of the regenerated plants will be capable of infection. This transformation is carried out by any of the techniques previously described.

A still further feature of the invention is a process for the production of an enzyme suitable for the stereospecific catalysis of an organic compound. The second nucleotide sequence of the chimeric nucleotide sequence comprises the transcribable sequence which leads to the production of the desired product. This process involves the infection of the appropriate host with a virus such as those described above, the growth of the infected host to produce the desired product and the isolation of the desired product. The growth of the infected host is in accordance with conventional techniques, as is the isolation of the resultant product. The stereospecific enzyme is then utilized to

catalyze the desired reaction. One use of sterespecific enzymes is in the separation of racemate mixtures.

In one example, a suitable esterase or lipase coding sequence such as isolated from an appropriate microorganism is inserted adjacent the promoter of the viral coat protein of a nucleotide sequence derived from TMV, oat mosaic virus (OMV) or rice necrosis virus (RNV) in which the coat protein coding sequence has been removed. A virus is prepared as described above using the resulting chimeric nucleotide sequence. Tobacco or germinating barley is infected with the appropriate virus. The chimeric nucleotide sequence self-replicates in the plant tissue to produce the esterase or lipase enzyme. This enzyme is isolated and used in the stereospecific preparation of a compound of formula (I), as described in EPA 0 233 656 or EPA 0 227 078.

An esterase coding sequence is isolated from the appropriate microorganism, such as *Bacillus subtilis*, *Bacillus licheniformis* (a sample of this species is deposited with the American Type Culture Collection, Rockville, Maryland (ATCC) under accession number 11945), *Pseudomonas fluorescens*, *Pseudomonas putida* (a sample of this species is deposited with the Institute for Fermentation (IFO), Osaka, Japan, under accession number 12996), *Pseudomonas riboflavina* (a sample of this species is deposited with IFO under accession number 13584), *Pseudomonas ovalis* (a sample of this species is deposited with the Institute of Applied Microbiology (IAM), University of Tokyo, Japan, under accession number 1049), *Pseudomonas aeruginosa* (IFO 13130), *Mucor angulimacrosporus* (IAM 6149), *Arthrobacterparaffineus* (ATCC 21218), *Strain is III-25* (CBS 666.86), *Strain LK 3-4* (CBS 667.86), *Strain Sp 4* (CBS 668.86), *Strain Thai III 18-1* (CBS 669.86), and *Strain Thai VI 12* (CBS 670.86). Advantageously, cultures of species *Bacillus subtilis* include cultures of species *Bacillus species*

Thai 1-8 (CBS 679.85), species *Bacillus species In IV-8* (CBS 680.85), species *Bacillus species Nap 10-M* (CBS 805.85), species *Bacillus species Sp III-4* (CBS 806.85), *Bacillus subtilis 1-85* (Yuki, S. et al., *Jpn.J.Genet. 42*, 251 (1967)), *Bacillus subtilis 1-85/pNAPT-7* (CBS 673.86), *Bacillus subtilis 1A-40/pNAPT-8* (CBS 674.86), and *Bacillus subtilis 1A-40/pNAPT-7* (CBS 675.86). Advantageously, cultures of *Pseudomonas fluorescens* include a culture of species *Pseudomonas species Kpr 1-6* (CBS 807.85), and *Pseudomonas fluorescens* species (IFO 3081.

A lipase coding sequence is isolated from the appropriate microorganism such as the genera *Candida, Rhizopus, Mucor, Aspergilus, Penicillium, Pseudomonas, Chromobacterium,* and *Geotrichium.* Particularly preferred is the lipase of *Candida cylindracea* (Qu-Ming et al., *Tetrahedron Letts. 27*, 7 (1986)).

EXAMPLES

The following examples further illustrate the present invention. In these examples, enzyme reactions were conducted in accordance with the manufacturer's recommended procedures unless otherwise indicated. Standard techniques such as those described in Maniatis, T. et al., *supra; Meth.in Enzymol.* Vol. 68, 100, 101, 118, 152-155, *supra;* and *DNA Cloning,* Vol. I, II, III, *supra,* were utilized for vector constructions and transformation unless otherwise specified.

EXAMPLE 1

Preparation of Chimeric TMV DNA Comprising Tyrosinase Gene in Place of the Coat Protein Gene

The 1.2 kb tyrosinase gene was derived from the pIJ702 streptomyces plasmid (available to the public) using a SacI/PVUII cut. It was inserted into the

SacI/EcoRV site of the Stratagene bluescript vector (KS), which is 2.9 kb in size. The resulting plasmid was maintained as pBG110 (4.1 kb). The tyrosinase gene was removed from pBG110 by a SacI/HindIII cut and inserted into pUC19 (2.7 kb) giving pBG115 (3.9 kb). The tyrosinase gene was removed from pBG115 by a EcoRI/HindIII cut and inserted into the Stratagene bluescript vector (KS+; 2.9 kb) giving pBG120 (4.1 kb). The tyrosinase gene was removed from pBG120 with a XhoI/SmaI cut and XhoI linkers were added to the SmaI end. This tyrosinase gene with XhoI ends was inserted into the XhoI site of the Stratagene bluescript vector (KS+; 2.9 kb) to give pBG130 (4.1 kb). Since there is an unwanted PstI site carried over in a poly-linker region, the pBG130 was modified by removing some of the non-coding bases and the poly-linker region. This was performed using a HindIII/SacI digest to release the tyrosinase from pBG130, then treating the 1.2 kb fragment with ExoIII to digest approximately 200 bases from the HindIII 3' end. After blunting the ends with T_4 DNA Polymerase, XhoI linkers were added and the 1.0 kb fragment was inserted into the Stratagene bluescript vector (Ks+; 2.9 kb) giving pBG132. The 1.0 kb tyrosinase gene was inserted into p803 (6.3 kb), which is a 3.6 kb subclone of the TMV plasmid, pS3-28 (available from W. O. Dawson, University of California, Riverside), in pUC19, giving plasmids pBG21 and pBG22 (depending on the direction of the tyrosinase gene; 7.3 kb each). The plasmid S3-28 (11.1 kb) is a clone of TMV that has had the coat protein gene removed with an XhoI site at that position. Dawson, W. O. et al., Phytopathology 78, 783 (1988). A NcoI/EcoRV cut of the pBG21 and pBG22 released a 2.4 kb fragment containing the tyrosinase gene. This fragment replaced a 1.4 kb piece in pS3-28 that had been removed. The resulting plasmids, pBG23 and pBG24 (depending on the direction of

the tyrosinase gene; 12.1 kb each) were the final DNA constructions that were used to make infectious RNA that were introduced into tobacco plants to introduce a mRNA for tyrosinase. This mRNA can be detected using northern blot procedures.

EXAMPLE 2

Preparation of Chimeric TMV DNA Comprising GUS Gene in Place of the Coat Protein Gene

The 1.8 kb GUS gene was derived from the pRAJ275 (4.5 kb; Clonetech Laboratories) using a EcoRI/NcoI cut. The sticky ends were filled in using the Klenow fragment. XhoI linkers were added and the fragment was inserted into the Stratagene bluescript vector (KS+), which is 2.9 kb in size. The resulting plasmid was maintained as pBG150 (4.7 kb). Using the techniques of Example 1, this GUS gene with the XhoI linkers was moved into the p803, giving pBG25 and pBG26 (depending on the direction of the GUS gene; 8.1 kb each). A SalI/NcoI cut of the pBG25 and pBG26 released a 3.8 kb fragment containing the GUS gene. These fragments replaced a 2.0 kb piece in pS3-28 that had been removed. The resulting plasmids, pBG27 and pBG28 (depending on the direction of the GUS gene; 12.9 kb each) were the final DNA constructions that were used to make infectious RNA that were introduced into tobacco plants to introduce a mRNA for the GUS gene. This mRNA can be detected using northern blot procedures. The activity of the GUS enzyme is also detectable.

EXAMPLE 3

Preparation of Chimeric TMV DNA Comprising GUS/Coat Protein Gene in Place of the TMV Coat Protein Gene

The 1.8 kb GUS gene from pBG150 (4.7 kbk), described in Example 2 above, was used for this construction. A PstI/NcoI cut of p35-5 released a 0.7 kb fragment that contains a portion of the 3' end and a portion of the 5' end of the TMV coat protein gene. The plasmid p35-5 (11.3 kb) is a clone of TMV that has had most of the coat protein gene removed and contains an XhoI site at the site where the internal portion of the coat protein gene is removed. Dawson, W. O. et al., Phytopathology 78, 783 (1988). This 0.7 kb fragment replaced a 0.5 kb piece in p803 that had been removed, giving pBG29 (6.5 kb). The 1.8 kb GUS gene was inserted into the XhoI site of pBG29, giving pBG31 and pBG32 (depending on the direction of the GUS/coat protein fusion gene; 8.3 kb each). A SalI/NcoI cut of the pBG31 and pBG32 released a 4.0 kb fragment containing the GUS/coat protein fusion gene. These fragments replaced a 2.0 kb piece in pS3-28 that had been removed. The resulting plasmids, pBG33 and pBG34 (depending on the direction of the GUS/coat protein fusion gene; 13.1 kb each) were the final DNA constructions that were used to make infectious RNA. pBG33, pBG34 and pS3-28 (as control) are transcribed into RNA which is used to infect tobacco plants by rubbing on the plants the RNA, an abrasive and a RNase inhibitor. Spots are noted on the plants indicating a successful infection. The infected plant tissue is macerated and fluorescence examined. Tissue infected with pS3-28 did not fluoresce whereas tissue infected with either pBG33 or pBG34 did fluoresce.

EXAMPLE 4

Preparation f a Non-Transmissible TMV Nucleotide Sequence

A full-length DNA copy of the TMV genome is prepared and inserted into the *Pst*I site of pBR322 as described by Dawson, W.O. et al., *Proc.Nat.Acad.Sci. USA 83*, 1832 (1986). The viral coat protein gene is located at position 5711 of the TMV genome adjacent the 30k protein gene. The vector containing the DNA copy of the TMV genome is digested with the appropriate restriction enzymes and exonucleases to delete the coat protein coding sequence. For example, the coat protein coding sequence is removed by a partial digestion with *Cla*I and *Nsi*I, followed by religation to reattach the 3'-tail of the virus. Alternatively, the vector is cut at the 3' end of the viral nucleic acid. The viral DNA is removed by digestion with *Bal*31 or exonuclease III up through the start codon of the coat protein coding sequence. A synthetic DNA sequence containing the sequence of the viral 3'-tail is then ligated to the remaining 5'-end. The deletion of the coding sequence for the viral coat protein is confirmed by preparing TMV RNA *in vitro* by using the lambda RNA polymerase promoter which was added to the 5' end of the TMV sequence in accordance with the technique described in Dawson, W.O. et al., *supra*. This material is used to infect tobacco plants. The isolated TMV RNA is found to be non-infective under natural conditions.

EXAMPLE 5

Preparation of a Non-Transmissible OMV Nucleotide Sequence

A full-length DNA copy of the OMV genome is prepared as described by Dawson, W.O. et al., *supra*. The vector containing the DNA copy of the OMV genome is

digested with the appropriate restriction enzymes or suitable exonucleases such as described in Example 1 to delete the coat protein coding sequence. The deletion of the coding sequence for the viral coat protein is confirmed by isolating OMV RNA and using it to infect germinating barley plants. The isolated OMV RNA is found to be non-infective under natural conditions.

EXAMPLE 6

Preparation of a Non-transmissible RNV Nucleotide Sequence

A full-length DNA copy of the RNV genome is prepared as described by Dawson, W.O. et al., *supra*. The vector containing the DNA copy of the RNV genome is digested with the appropriate restriction enzymes or suitable exonucleases such as described in Example 1 to delete the coat protein coding sequence. The deletion of the coding sequence for the viral coat protein is confirmed by isolating RNV RNA and using it to infect germinating barley plants. The isolated RNV RNA is found to be non-infective under natural conditions.

EXAMPLE 7

Preparation of Nucleotide Sequence Containing an Esterase Coding Sequence from *Bacillus*

The coding sequence for an esterase is isolated from *Bacillus subtilis Thai 1-8* (CBS 679.85) as follows.

The positive selection vector pUN121 (Nilsson et al., *Nucleic Acids Res. 11*, 8019 (1983)) is used. This vector carries an ampicillin resistance gene, a tetracyclin resistance gene and a C_1-repressor gene. Transcription of the tetracyclin gene is prevented by the gene product of the C_1-repressor gene. Insertion of foreign DNA into the *BclI* site of the C_1-repressor gene

results in activation of the tetracycline gene. This allows a positive selection of r combinants on ampicillin/tetracyclin agar plates.

Partially *Sau*3a-digested *Bacillus subtilis Thai 1-8* DNA is mixed with *Bcl*I-digested pUN121 DNA. After recirculation by the use of polynucleotide ligase, the DNA mixture is introduced into *E. coli* DH1 (ATCC 33849) using the $CaCl_2$ transformation procedure. One thousand *E. coli* colonies are obtained which are resistant to ampicillin and tetracycline. All transformants are stored and replica-plated according to Gergan et al., *Nucl.Acids Res. 7*, 2115 (1979). Replicated colonies are screened using a soft agar overlay technique based on a previously described procedure to detect esterase activity (Higerd et al., *J.Bacteriol. 114*, 1184 (1973)). Essentially a mixture of 0.5% low-melting agarose, 0.5 M potassium phosphate (pH 7.5), 0.5 mg/l β-naphthyl acetate and 0.5 mg/ml fast-blue is spread over the transformants. Within a few minutes colonies with esterase or lipase activity develop purple color. Such colonies are grown overnight in 2* YT (16 g/l Bactotryptone, 10 g/l Yeast Extract, 5 g/l NaCl) medium and subsequently assayed for their ability to convert S-naproxen ester to S-naproxen (the method of Example 1 of EPA 0 233 656). One *E. coli* transformant is able to convert S-naproxen ester. The plasmid isolated from this transformant, which is called pNAPT-2 (CBS 67186), is shown in Figure 1. Its size is 9.4 kb.

*Hind*III restriction enzyme fragments of pNAPT-2 are ligated into pPNEO/ori. This is performed as described below. pPNeo/ori is constructed by ligating the 2.7 kb *Eco*RI-*Sma*I restriction fragment of pUC19 (Yanish et al., *Gene 33*, 103 (1985)) to the 2.5 kb *Eco*RI-*Sna*BI restriction fragment of pUB110 (Gryczan et al., *J.Bacteriol. 134*, 318 (1978)). The resulting shuttle plasmid, pPNeo/ori (5.2 kb) has the capacity to replicate both in

E. coli and in *Bacillus* species due to the pr sence of the pUC19 origin, and the pUB110 origin. In addition, pPNeo/ori carries a gene encoding ampicillin resistance and a gene encoding neomycin resistance (Yanish-Perron et al., *supra*; Matsumura et al., *J.Bacteriol. 160*, 413 (1984)).

For subcloning, *Hind*III-digested pNAPT-2 is mixed with *Hind*III-digested pPNeo/ori and ligated. The mixture is transformed to *E. coli* JM101 hsds as described (Maniatis et al., *supra*). *E. coli* JM101 hsds is obtained from the Phabagen collection (accession number PC 2493, Utrecht, The Netherlands). Colonies capable of hydrolyzing ß-naphthyl acetate are selected as described in Example 10 of EPA 0 233 656. From two positive colonies, plasmid DNA is isolated and characterized in detail by determining several restriction enzyme recognition positions. The physical maps of these plasmids, pNAPT-7 and pNAPT-8, are shown in Figures 2 and 3, respectively.

EXAMPLE 8

Preparation of Nucleotide Sequence Containing an Esterase Coding Sequence from *E. Coli*

Plasmids pIP1100 (isolated from *E. coli* BM 2195) and pBR322 are mixed, digested with *Ava*I, ligated and transformed into *E. coli*, and clones are selected on Em (200 µg/ml). Transformants resistant to Ap and Em but also to Sm are analyzed by agarose gel electrophoresis of crude lysates (Birnboim, H.C. et al., *Nucl.Acids Res.* 7, 1513 (1979). The transformant harboring the smallest hybrid plasmid is selected, its plasmid DNA is digested with *Ava*I, and the 3.5 kb pIP1100 insert is purified and partially digested with *Sau*3A. The restriction fragments obtained are cloned into the *Bam*HI site of pBR322 and transformants selected on Em are replica-

plat d on Sm. The plasmid content of transformants resistant only to Ap and Em is analyzed by agarose gel electrophoresis. DNA from the smallest hybrid, pAT63, is purified and analyzed by agarose gel electrophoresis after digestions with *Sau*3A, *Eco*RI, *Pst*I or *Hind*III-*Bam*HI endonucleases (not shown). Plasmid pAT63 consists of pBR322 plus a 1.66 kb pIP1100 DNA insert. Purified *Eco*RI-*Hind*III (1750-bp) and *Bam*HI-*Pst*I (970-bp) fragments of pAT63 are subcloned into pUC8 and found not to confer resistance to Em.

The *Hpa*II-*Bam*HI fragment of pAT63 is sequenced according to Sanger, F. et al., *Proc. Nat. Acad. Sci. USA 74*, 5463 (1977). The nucleotide sequence of the 5' end of the fragment is shown in Figure 4. The complete sequence is shown in Ounissi, H. et al., *Gene 35*, 271 (1985).

EXAMPLE 9

Preparation of Virus Containing Esterase from *Bacillus*

The esterase gene described in Example 7 is removed from the vector, blunt-ended using Mung Bean nuclease or DNA polymerase I, and *Xho*I linkers added. This esterase with *Xho*I linkers is cleaved with *Xho*I and inserted into the coat protein site of the "non-infective nucleotide sequence" described in Examples 4, 5 or 6. This new chimeric plasmid contains pBR322 (which can be removed with *Pst*I restriction endonuclease) followed by the lambda RNA polymerase promoter, and the viral, e.g., TMV, genome with the esterase gene inserted at the *Xho*I site. A plasmid identified as pBGC001, (with the TMV genome and the *Bacillus* esterase gene) as shown in Figure 5, is isolated. A CaMV 35S promoter containing nucleotides -168 to +9 and having a *Cla*I restriction site at the 5' end and a *Hind*III restriction site at the 3' end is prepared as described by Odell, J.T. et al.,

Nature 313, 810 (1985). The 35S promoter is isolated by cleaving with *Cla*I and *Hind*III and separating the 35S promoter containing the *Cla*I/*Hind*III fragment. The *Cla*I*Hin*dIII fragment is blunt-ended using Mung Bean nuclease or DNA polymerase I, *Pst*I linkers added and cleaved with *Pst*I. The plasmid pBGC001 is partially digested with *Pst*I and ligated with the 35S promoter fragment with *Pst*I linkers. A plasmid identified as pBGC002 is identified and isolated which contains an intact promoter including its ribosome binding site at the 35S promoter located at the *Pst*I site at position 11493 of pBGC001. Viral chimeric RNA is prepared *in vitro* as described by Dawson, W.O. et al., *supra*. The viral chimeric RNA is inoculated into production plants or plant cells and large volumes of replicate copies of the viral chimeric RNA are made.

A second vector is prepared by inserting the viral coat protein coding sequence, isolated in Examples 4, 5 or 6, adjacent the *lac* promoter in the vector pBR322. This vector is used to transform *E. coli* as the production cells. The production cells are grown and the resultant viruses are isolated. The coat protein and replicate copies of the viral chimeric RNA are then combined to form the virus.

EXAMPLE 10

Production of Esterase from *Bacillus*

The viruses isolated in Example 9 are used to infect tobacco plants (viruses based on TMV) or germinating barley plants (viruses based on OMV or RNV). The infected plants are grown under normal growth conditions. The plants produce esterase which is isolated by conventional techniques.

EXAMPLE 11

Preparation of Virus Containing Esterase from *E. coli*

The esterase gene described in Example 8 is removed from the vector, blunt-ended using Mung Bean nuclease or DNA polymerase I, and *Xho*I linkers added. This esterase with *Xho*I linkers is cleaved with *Xho*I and inserted into the coat protein site of the "non-infective nucleotide sequence" described in Examples 4, 5 or 6. This new chimeric plasmid contains pBR322 (which can be removed with *Pst*I restriction endonuclease) followed by the lambda RNA polymerase promoter, and the viral, e.g. TMV, genome with the esterase gene inserted at the *Xho*I site. A plasmid identified as pBGC001-1 (with TMV genome and the *E. coli* esterase gene), similar to that shown in Figure 5, is isolated. A CaMV 35S promoter containing nucleotides -168 to +9 and having a *Cla*I restriction site at the 5' end and a *Hin*dIII restriction site at the 3' end is prepared as described by Odell, J.T. et al., *Nature 313*, 810 (1985). The 35S promoter is isolated by cleaving with *Cla*I and *Hin*dIII and separating the 35S promoter containing the *Cla*I/*Hin*dIII fragment. The *Cla*I*Hin*dIII fragment is blunt-ended using Mung Bean nuclease or DNA polymerase I, *Pst*I linkers added and cleaved with *Pst*I. The plasmid pBGC001-1 is partially digested with *Pst*I and ligated with the 35S promoter fragment with *Pst*I linkers. A plasmid identified as pBGC002-1 is identified and isolated which contains an intact promoter including its ribosome binding site at the 35S promoter located at the *Pst*I site at position 11493 of pBGC001-1. Viral chimeric RNA is prepared *in vitro* as described by Dawson, W.O. et al., *supra*. The viral chimeric RNA is inoculated into production plants or plant cells and large volumes of replicate copies of the viral chimeric RNA are made.

A second vector is prepared by inserting the viral coat protein coding sequence, isolated in Exampl s 4, 5 or 6, adjacent the *lac* promoter in the vector pBR322. This vector is used to transform the *E. coli* as the production cells. The production cells are grown and the resultant viruses are isolated. The coat protein and replicate copies of the viral chimeric RNA are then combined to form the virus.

EXAMPLE 12

Production of Esterase from *E. coli*

The viruses isolated in Example 11 are used to infect tobacco plants (viruses based on TMV) or germinating barley plants (viruses based on OMV or RNV). The infected plants are grown under normal growth conditions. The plants produce esterase which is isolated by conventional techniques.

EXAMPLE 13

Preparation of Nucleotide Sequence Containing An Acylase Coding Sequence From *Arthrobacter Viscosus*

The coding sequence from acylase is isolated from *Arthrobacter viscosus* 8895GU, ATCC 27277 follows.

A gene library of *A. Viscosus* 8895GU is constructed by inserting *Eco*RI-cleaved *A. viscosus* chromosomal DNA into the *Eco*RI cleavage site of pACYC184. Chang, A. C.Y. et al., *J. Bacteriol. 134*, 1141 (1978). The vector DNA and *A. viscosus* DNA are both digested with *Eco*RI. The 5' end of the vector DNA is dephosphorylated with calf intestinal alkaline phosphatase. Dephosphoroylated vector DNA and digested *A. viscosus* DNA are incubated with T4 DNA ligase and transformed into *E. coli* HB101. Transformed colonies of *E. coli* were screened by the *Serratia marcescens* overlay technique. Penicillin G was

added to the medium. *S. marcescens* is sensitive to the deacylation product of penicillin G, 6-aminopenicillamic acid (6-APA). Colonies of transformed *E. coli* will produce areas of *S. marcescens* inhibition in overnight cultures. The plasmid carried by transformed *E. Coli* is referred to as pHYM-1. The plasmid having the opposite DNA orientation is designated pHYM-2. Ohashi, H. et al., *Appl. Env. Microbiol. 54 (11)*, 2603 (1988).

EXAMPLE 14

Preparation of Virus Containing Acylase from *Arthrobacter*

The acylase gene described in Example 13 is removed from the vector by digestion with *Eco*RI. The gene is blunt-ended using Mung Bean nuclease or DNA polymerase I, and *Xho*I linkers added. This acylase gene with *Xho*I linkers is partially digested with *Xho*I and inserted into the coat protein site of the "non-infective nucleotide sequence" described in Examples 4, 5 or 6. This new chimeric plasmid contains pBR322 (which can be removed with *Pst*I restriction endonuclease) followed by the lambda RNA polymerase promoter, and the viral, e.g. TMV, genome with the acylase gene inserted at the *Xho*I site. A plasmid identified as pBGC013, is isolated. It is similar to the pBGC001, shown in Figure 5, except that the acylase gene is in place of the esterase gene.

A CaMV 35S promoter containing nucleotides -168 to +9 and having a *Cla*I restriction site at the 5' end and a *Hin*dIII restriction site at the 3' end is prepared as described by Odell, J.T. et al., *Nature 313*, 810 (1985). The 35S promoter is isolated by cleaving with *Cla*I and *Hin*dIII and separating the 35S promoter containing the *Cla*I/*Hin*dIII fragment. The *Cla*I*Hin*dIII fragment is blunt-ended using Mung Bean nuclease or DNA polymerase I, *Pst*I linkers added and cleaved with *Pst*I. The plasmid pBGC013 is partially digested with *Pst*I and

ligated with the 35S promoter fragment with *Pst*I linkers. A plasmid identified as pBGC014 is identified and isolated which contains an intact promoter including its ribosome binding site at the 35S promoter located at the *Pst*I site of pBGC013. Viral chimeric RNA is prepared *in vitro* as described by Dawson, W.O. et al., *supra*. The viral chimeric RNA is inoculated into production plants or plant cells and large volumes of replicate copies of the viral chimeric RNA are made.

A second vector is prepared by inserting the viral coat protein coding sequence, isolated in Examples 4, 5 or 6, adjacent the *lac* promoter in the vector pBR322. This vector is used to transform *E. coli* as the production cells. The production cells are grown and the resultant viruses are isolated. The coat protein and replicate copies of the viral chimeric RNA are then combined to form the virus.

EXAMPLE 15

Production of Acylase from *Arthrobacter*

The viruses isolated in Example 14 are used to infect tobacco plants (viruses based on TMV) or germinating barley plants (viruses based on OMV or RNV). The infected plants are grown under normal growth conditions. The plants produce acylase which is isolated by conventional techniques.

EXAMPLE 16

Preparation of a Non-Transmissible Eastern Equine Encephalomyelitis Virus Nucleotide Sequence

A full-length cDNA copy of the Eastern Equine Encephalomyelitis Virus (EEEV) genome is prepared and inserted into the *Pst*I site of pUC18 as described by Chang, G-J. J. et al., *J. Gen. Virol.*, *68*, 2129 (1987).

The sequence for the viral coat protein and its adjacent E1 and E2 glycoprot in transmissibility factors are located on the region corresponding to the 26S RNA region. The vector containing the cDNA copy of the EEEV genome is digested with the appropriate restriction enzymes and exonucleases to delete the coding sequence of the coat protein and the E1 and E2 proteins (structural protein coding sequence).

For example, the structural protein coding sequence is removed by partial digestion with *Mbo*I, followed by religation to remove a vital portion of the structural gene. Alternatively, the vector is cut at the 3' end of the viral structural gene. The viral DNA is sequentially removed by digestion with *Bal*31 or Micrococcal S1 nuclease up through the start codon of the structural protein sequence. The DNA sequence containing the sequence of the viral 3'-tail is then ligated to the remaining 5'-end. The deletion of the coding sequence for the structural proteins is confirmed by isolating EEEV RNA and using it to infect an equine cell culture. The isolated EEEV RNA is found to be non-infective under natural conditions.

Alternatively only the coding sequence for the coat protein is deleted and the sequence for the E1 and E2 glycoproteins remain in the vector containing the cDNA copy of the EEEV genome. In this case, the coat protein coding sequence is removed by partial digestion with *Mbo*I followed by religation to reattach the 3'-tail of the virus. This will remove a vital portion of the coat protein gene.

A second alternative method for removing only the coat protein sequence is to cut the vector at the 3'-end of the viral coat protein gene. The viral DNA is removed by digestion with *Bal*31 or Micrococcal S1 nuclease up through the start codon of the coat protein sequence.

The synthetic DNA sequenc containing the sequence of the 3'-tail is then ligated to the remaining 5'-end.

The deletion of the coding sequence for the coat protein is confirmed by isolating EEEV RNA and using it to infect an equine cell culture. The isolated EEEV RNA is found to be non-infective under natural conditions.

EXAMPLE 17

Preparation of Nucleotide Sequence Containing A Lipase Coding Sequence From Human Gastric Tissue

mRNA is prepared by guanidinium isothiocyanate extraction of frozen tissue. Polyadenylated RNA is isolated by oligo(dT)-cellulose chromatography. cDNA is prepared from human stomach mRNA by procedures well known in the art. cDNA is annealed to *Pst*I cut dG-tailed pBR322. The hybrid plasmid is transformed into *E. coli* DH1. Transformants are screened by colony hybridization on nitrocellulose filters. The probe used is synthesized from the rat lingual lipase gene and labeled by Nick translation. Positive colonies are grown up and plasmids are analyzed by restriction endonuclease mapping.

EXAMPLE 18

Preparation of Virus Containing Human Gastric Lipase

The lipase gene described in Example 17 is removed from the vector, blunt-ended using Mung Bean nuclease or DNA polymerase I, and *Mbo*I linkers added. This lipase gene with *Mbo*I linkers is cleaved with *Mbo*I and inserted into the coat protein site of the "non-infective EEEV nucleotide sequence" described in Example 16. This new chimeric plasmid contains pUC18 (which can be removed with *Pst*I restriction endonuclease), and the EEEV genome

with th lipase gene inserted at the *Mbo*I site. A plasmid identified as pBGC015, is isolated. A CaMV 35S promoter containing nucleotides -168 to +9 and having a *Cla*I restriction site at the 5' end and a *Hin*dIII restriction site at the 3' end is prepared as described by Odell, J.T. et al., *Nature 313*, 810 (1985). The 35S promoter is isolated by cleaving with *Cla*I and *Hin*dIII and separating the 35S promoter containing the *Cla*I/*Hin*dIII fragment. The *Cla*I/*Hin*dIII fragment is blunt-ended using Mung Bean nuclease or DNA polymerase I, *Pst*I linkers added and cleaved with *Pst*I. The plasmid pBGC015 is partially digested with *Pst*I and ligated with the 35S promoter fragment with *Pst*I linkers. A plasmid identified as pBGC016 is identified and isolated which contains an intact promoter including its ribosome binding site at the 35S promoter located at the *Pst*I site at position 11493 of pBGC013. Viral chimeric RNA is prepared *in vitro* as described by Dawson, W.O. et al., *supra*. The viral chimeric RNA is inoculated into production plants or plant cells and large volumes of replicate copies of the viral chimeric RNA are made.

A second vector is prepared by inserting the viral structural protein coding sequence, isolated in Example 16 with *XhoI* linkers add, adjacent the *lac* promoter in the vector pBR322. This vector is used to transform *E. coli* as the production cells. The production cells are grown and the resultant viruses are isolated. The coat protein and replicate copies of the viral chimeric RNA are then combined to form the virus.

EXAMPLE 19

Preparation of Lipase *In Vitro*

The viruses isolated in Example 18, made by the combination of chimeric nucleotide sequence and coat

protein are used to infect equine cell cultures (viruses based on EEEV). The infected cell cultures are grown under normal cell culture growth conditions. The cells produce lipase which reacts with substrates when in contact with the cells or after the lipase is extracted from the cells. The desired stereoisomer is isolated by conventional techniques.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications. This application is intended to cover any variations, uses or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known and customary practice within the art to which the invention pertains.

WHAT IS CLAIMED IS:

1. A chimeric nucleotide sequence comprising a first nucleotide sequence which has substantial sequence homology to a viral nucleotide sequence which is capable of replication and is biologically contained and a second nucleotide sequence which is a coding sequence for an enzyme and is capable of transcription in a host, said enzyme capable of enantiospecific catalysis of an organic compound.

2. The chimeric nucleotide sequence of claim 1 wherein said viral nucleotide sequence is a prokaryotic viral nucleotide sequence.

3. The chimeric nucleotide sequence of claim 1 wherein said viral nucleotide sequence is a eukaryotic viral nucleotide sequence.

4. The chimeric nucleotide sequence of claim 1 wherein said viral nucleotide sequence is a plant viral nucleotide sequence.

5. The chimeric nucleotide sequence of claim 1 wherein said chimeric nucleotide sequence is selected from the group consisting of DNA, RNA and cDNA.

6. The chimeric nucleotide sequence of claim 1 wherein said viral nucleotide sequence lacks a biologically functional viral coat protein coding sequence.

7. The chimeric nucleotide sequence of claim 6 wherein said viral nucleotide sequence further lacks a biologically functional viral transmissability factor coding sequence.

8. The chimeric nucl tide sequence of claim 6 wherein the transcription of said second nucleotide sequence is regulated by a viral coat protein promoter.

9. The chimeric nucleotide sequence of claim 1 wherein said enzyme is selected from the group consisting of an esterase, a lipase and a phosphatase.

10. A virus comprising a chimeric nucleotide sequence which comprises a first nucleotide sequence which has substantial sequence homology to a viral nucleotide sequence and which is capable of replication and is biologically contained and a second nucleotide sequence which is a coding sequence for an enzyme and is capable of transcription in a host, said enzyme capable of enantiospecific catalysis of an organic compound.

11. The virus of claim 10 wherein said viral nucleotide sequence is a prokaryotic viral nucleotide sequence.

12. The virus of claim 10 wherein said viral nucleotide sequence is a eukaryotic viral nucleotide sequence.

13. The virus of claim 10 wherein said viral nucleotide sequence is a plant viral nucleotide sequence.

14. The virus of claim 10 wherein said chimeric nucleotide sequence is selected from the group consisting of DNA, RNA and cDNA.

15. The virus of claim 10 wherein said viral nucleotide sequence lacks a biologically functional viral coat protein coding sequence.

16. Th virus of claim 15 wherein said viral nucleotide sequence further lacks a biologically functional viral transmissability factor coding sequence.

17. The virus of claim 15 wherein the transcription of said second nucleotide is regulated by a viral coat protein promoter.

18. The chimeric nucleotide sequence of claim 10 wherein said enzyme is selected from the group consisting of an esterase, a lipase and a phosphatase.

19. A vector comprising a chimeric nucleotide sequence adjacent a nucleotide sequence selected from the group consisting of a production cell promoter and an origin of replication compatible with said production cell, said chimeric nucleotide sequence comprising a first nucleotide sequence which has substantial sequence homology to a viral nucleotide sequence which is capable of replication and is biologically contained and a second nucleotide sequence which is a coding sequence for an enzyme and is capable of transcription in a host, said enzyme capable of enantiospecific catalysis of an organic compound.

20. The vector of claim 19 wherein said viral nucleotide sequence is a prokaryotic viral nucleotide sequence.

21. The vector of claim 19 wherein said viral nucleotide sequence is a eukaryotic viral nucleotide sequence.

22. The vector of claim 19 wherein said viral nucleotide sequence is a plant viral nucleotide sequence.

23. The vector of claim 19 wherein said chimeric nucleotide sequence is selected from the group consisting of DNA, RNA and cDNA.

24. The vector of claim 19 wherein said viral nucleotide sequence lacks a biologically funactional viral coat protein coding sequence.

25. The vector of claim 24 wherein said viral nucleotide sequence further lacks a biologically funactional viral transmissability factor coding sequence.

26. The vector of claim 24 wherein the transcription of said second nucleotide sequence is regulated by a viral coat protein promoter.

27. The vector of claim 19 wherein the chimeric nucleotide sequence of claim 1 wherein said enzyme is selected from the group consisting of an esterase, a lipase and a phosphatase.

28. A production cell having a vector which comprises a chimeric nucleotide sequence adjacent a nucleotide sequence selected from the group consisting of a production cell promoter and an origin of replication compatible with said production cell, said chimeric nucleotide sequence comprising a first nucleotide sequence which has substantial sequence homology to a viral nucleotide sequence which is capable of replication and is biologically contained and a second nucleotide

sequenc which is a coding sequ nce for an nzyme and is capable of transcripti n in a host, said enzyme capable of enantiospecific catalysis of an organic compound.

29. The production cell of claim 29 wherein said viral nucleotide sequence is a prokaryotic viral nucleotide sequence.

30. The production cell of claim 29 wherein said viral nucleotide sequence is a eukaryotic viral nucleotide sequence.

31. The production cell of claim 29 wherein said viral nucleotide sequence is a plant viral nucleotide sequence.

32. The production cell of claim 28 wherein said chimeric nucleotide sequence is selected from the group consisting of DNA, RNA and cDNA.

33. The production cell of claim 28 wherein said viral nucleotide sequence lacks a biologically functional viral coat protein coding sequence.

34. The production cell of claim 33 wherein said viral nucleotide sequence further lacks a biologically functional viral transmissability factor coding sequence.

35. The production cell of claim 33 wherein the transcription of said second nucleotide sequence is regulated by a viral coat protein promoter.

36. The production cell of claim 28 wherein the chimeric nucleotide sequence of claim 1 wherein

said enzyme is selected from the group consisting of an esterase, a lipas and a phosphatase.

37. The production cell of claims 28, 33 or 35 which further has a second vector which comprises a viral coat protein coding sequence adjacent a production cell promoter, said viral coat protein being compatible with said first nucleotide sequence and being biologically functional.

38. The production cell of claim 28, 33 or 35 which has been stably transformed to contain a coding sequence for a viral replicase useful for the replication of said chimeric nucleotide sequence.

39. The production cell of claim 28 or 33 which has been stably transformed to contain a coding sequence for a viral coat protein compatible with said chimeric nucleotide sequence.

40. The production cell of claim 39 which has been further stably transformed to contain a coding sequence for a viral replicase useful for the replication of said chimeric nucleotide sequence.

41. The production cell of claim 28 or 33 which has been further stably transformed to contain a coding sequence for viral transmissibility factor compatible with said chimeric nucleotide sequence.

42. The production cell of claim 41 which has been further stably transformed to contain a coding sequence for a replicase useful for the replication of said chimeric nucleotide sequence.

43. Th production cell of claim 28 or 33 which has been stably transformed to contain a coding sequence for viral coat protein and a viral transmissibility factor compatible with said chimeric nucleotide sequence.

44. The production cell of claim 28 or 33 which has been stably transformed to contain coding sequences for a replicase, a viral coat protein and a viral transmissibility fctor compatible with said chimeric nucleotide sequence.

45. The production cell of claim 35 which has been stably transformed to contain coding sequences for a replicase, a viral coat protein and a viral transmissibility factor compatible with said chimeric nucleotide sequence.

46. A host having a chimeric nucleotide sequence which comprises a first nucleotide sequence which has substantial sequence homology to a viral nucleotide sequence and which is capable of replication and is biologically contained and a second nucleotide sequence which is a coding sequence for an enzyme and is capable of transcription in said host, said enzyme capable of enantiospecific catalysis of an organic compound.

47. The host of claim 46 wherein said viral nucleotide sequence is a prokaryotic viral nucleotide sequence.

48. The host of claim 46 wherein said viral nucleotide sequence is a eukaryotic viral nucleotide sequence.

49. The host of claim 46 wherein said viral nucleotide sequence is a plant viral nucleotid sequ nce.

50. The host of claim 48 wherein said chimeric nucleotide sequence is selected from the group consisting of DNA, RNA and cDNA.

51. The host of claim 46 wherein said viral nucleotide sequence lacks a biologically functional viral coat protein coding sequence.

52. The host of claim 51 wherein said viral nucleotide sequence further lacks a biologically functional viral transmissability factor coding sequence.

53. The host of claim 51 wherein the transcription of said second nucleotide sequence is regulated by a viral coat protein promoter.

54. The host of claim 46 wherein the chimeric nucleotide sequence of claim 1 wherein said enzyme is selected from the group consisting of an esterase, a lipase and a phosphatase.

55. The host of claim 46 which has been stably transformed to contain a coding sequence for a replicase useful for the replication of said chimeric nucleotide sequence.

56. The host of claim 51 which has been stably transformed to contain a coding sequence for a replicase useful for the replication of said chimeric nucleotide sequence.

57. A process for producing an optically active organic compound which comprises (a) infecting a host with

a virus comprising a chimeric nucl otide sequ nce which comprises a first nucleotide s quence which has substantial sequence homology to a viral nucleotide sequence and which is capable of replication and is biologically contained and a second nucleotide sequence which is a coding sequence for an enzyme and is capable of transcription in a host, said enzyme capable of enantiospecific catalysis of an organic compound, (b) growing the infected host to produce said enzyme, (c) isolating said enzyme and (d) reacting said enzymes with the organic compound.

58. A process of producing an enantiospecific enzyme in a transgenic production organism or part thereof, said organism contains one or more coding sequences for (i) a viral replicase, (ii) a viral coat protein, (iii) a viral transmissibility factor or (iv) combinations thereof, said process comprises (a) infecting one or more cells of said organism with an infectious virus said virus cmoprising a chimeric nucleotide sequence which comprises a first nucleotide sequence which has substantial sequence homology to a viral nucleotide sequence and which is capable of replication and is biologically contained, and a second nucleotide sequence which is a coding sequence for said enzyme capable of transcription in a host and (b) growing the transgenic organism, whereby said transgenic organism produces said enzyme.

59. The process of claim 57 or 58 wherein said viral nucleotide sequence is a prokaryotic viral nucleotide sequence.

60. The process of claim 57 or 58 wherein said viral nucleotide sequence is a eukaryotic viral nucleotide sequence.

61. The process of claim 57 or 58 wherein said viral nucleotide sequence is a plant viral nucleotide sequence.

62. The process of claim 57 or 58 wherein said chimeric nucleotide sequence is selected from the group consisting of DNA, RNA and cDNA.

63. The process of claim 57 or 58 wherein said viral nucleotide sequence lacks a viral coat protein coding sequence.

64. The process of claim 63 wherein said viral nucleotide sequence further lacks a viral transmissability factor coding sequence.

65. The process of claim 57 or 58 wherein the transcription of said second nucleotide sequence is regulated by a viral coat protein promoter.

66. The process of claim 57 or 58 wherein the chimeric nucleotide sequence of claim 1 wherein said enzyme is selected from the group consisting of an esterase, a lipase and a phosphatase.

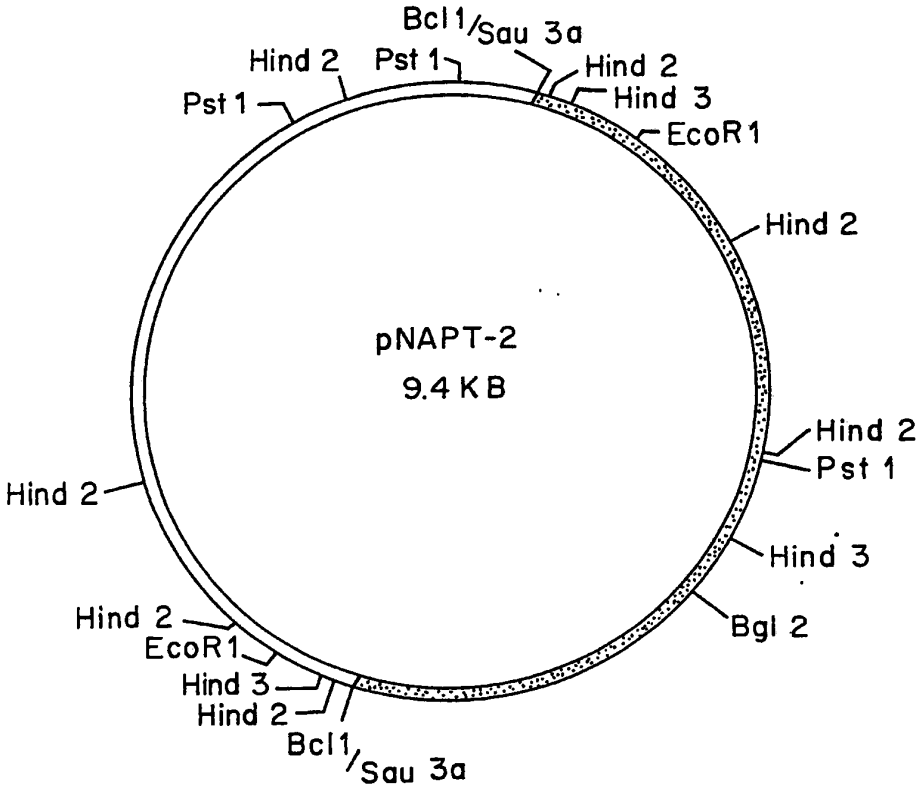

Bcl 1/Sau 3a
Hind 2
Pst 1
Hind 2
Hind 3
EcoR1
Pst 1
Hind 2
pNAPT-2
9.4 KB
Hind 2
Pst 1
Hind 2
Hind 3
Bgl 2
Hind 2
EcoR1
Hind 3
Hind 2
Bcl 1/Sau 3a

FIG. 1

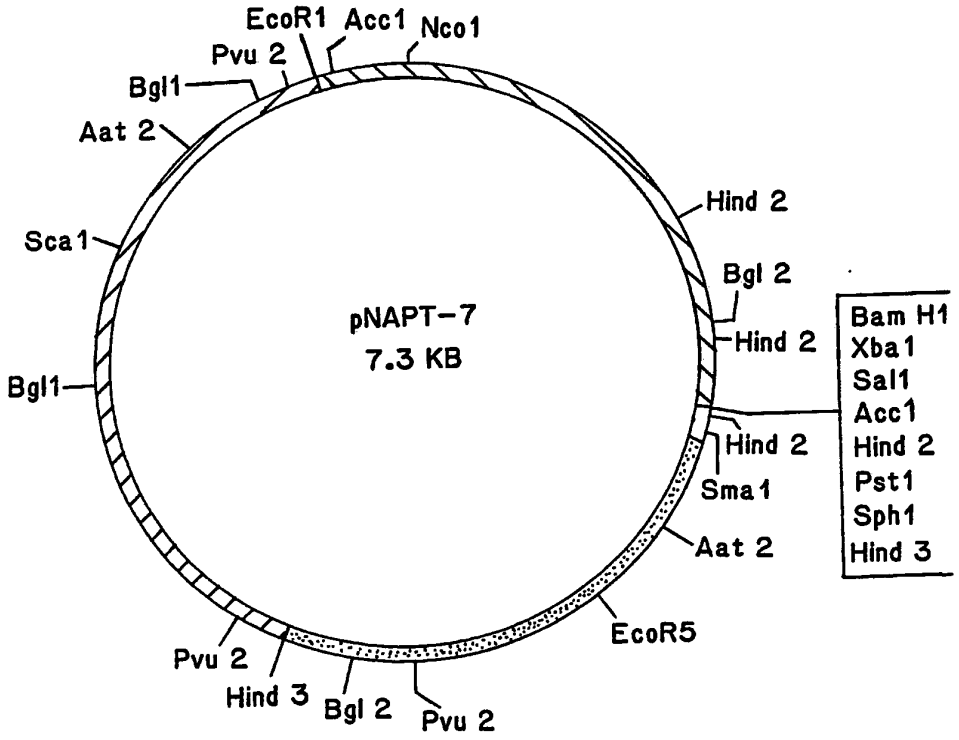

EcoR1
Acc1
Nco1
Pvu 2
Bgl1
Aat 2
Sca1
Bgl1
pNAPT-7
7.3 KB
Hind 2
Bgl 2
Hind 2
Hind 2
Sma1
Aat 2
EcoR5
Pvu 2
Hind 3
Bgl 2
Pvu 2
Bam H1
Xba1
Sal1
Acc1
Hind 2
Pst1
Sph1
Hind 3

FIG. 2

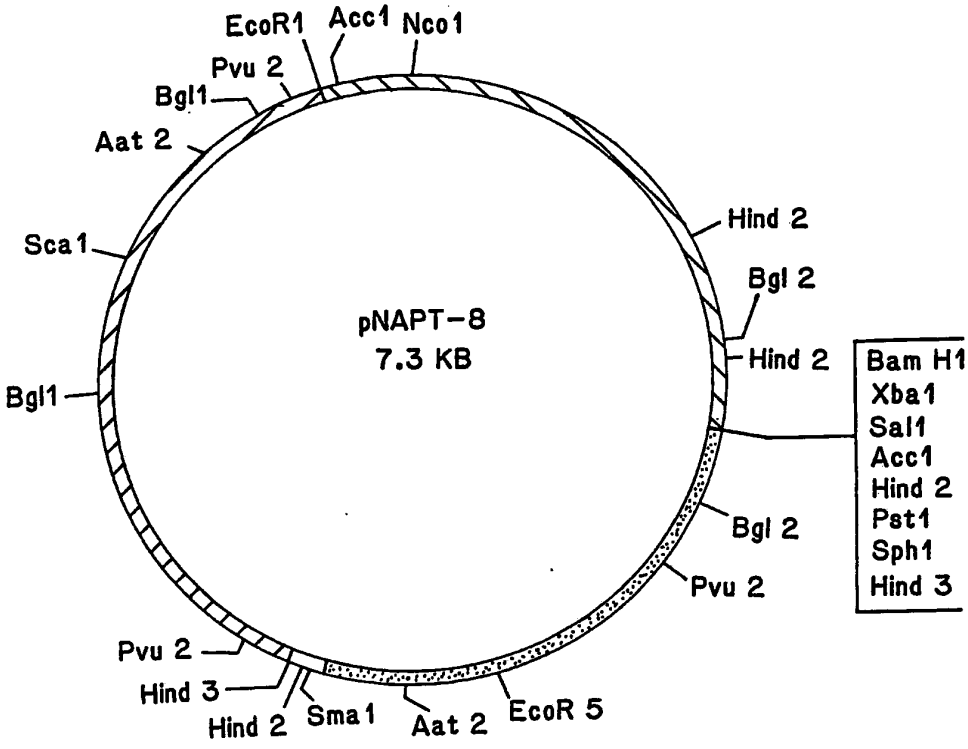

EcoR1
Acc1
Nco1
Pvu 2
Bgl1
Aat 2
Sca1
Bgl1
pNAPT-8
7.3 KB
Hind 2
Bgl 2
Hind 2
Bgl 2
Pvu 2
Pvu 2
Hind 3
Hind 2
Sma1
Aat 2
EcoR 5
Bam H1
Xba1
Sal1
Acc1
Hind 2
Pst1
Sph1
Hind 3

FIG. 3

```
CCGGGTTATT GGCGAACGAA GTATGACGAT TACAGCAATA AACGCAAAGG TAAAAAAATG      60

ACATTGGAGA ACGACCAGAA CACTTTTACA GCCTCAAAAG CTGGACTTCA ATGAGTTTGA     120

                                    -35                      -10
GATTCTTACT TCCGTAATTG AGGGCSCCCG AATTGTCGGC ATTGGCGAGG GCGCTCATTT     180

                                                             RBS
TGTCGCGGAG TTTTCACTGG CTAGAGCTAG TCTTATCCGC TATTTGGTCG AAAGGCATGA     240

        MET ARG LEU VAL TRP LYS CYS GLY
GTTTAA ATG CGA TTG GTT TGG AAA TGT GGG                                 300
```

FIG. 4

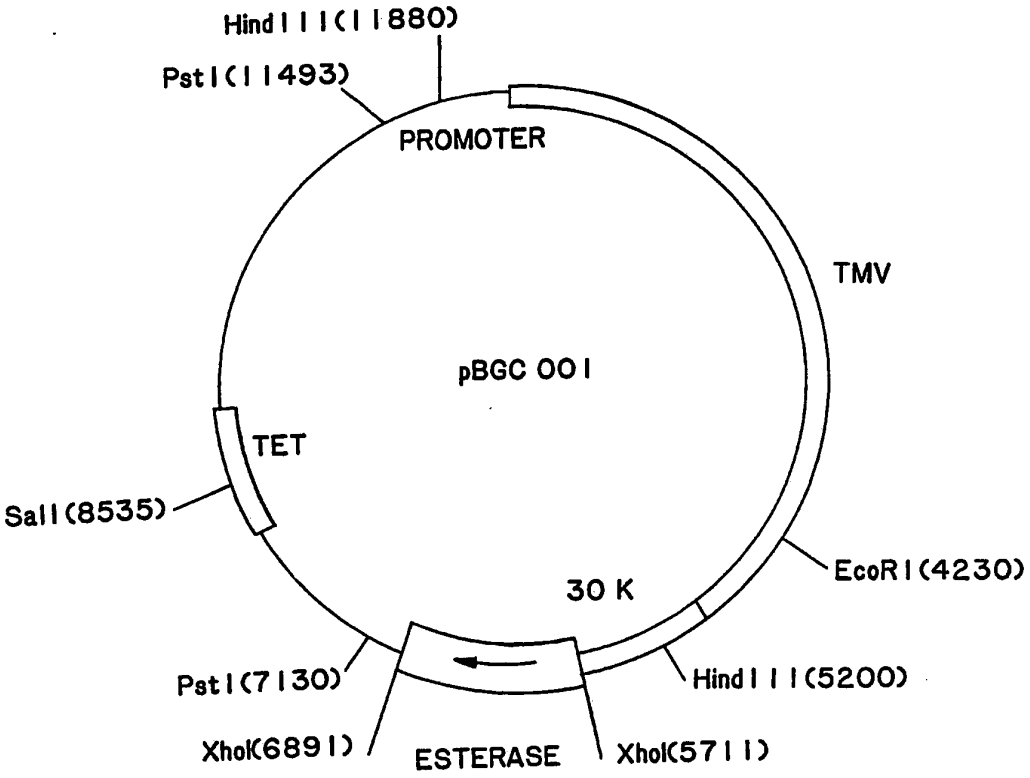

Hind III (11880)
Pst I (11493)
PROMOTER
TMV
pBGC 001
TET
Sal I (8535)
EcoR I (4230)
30 K
Pst I (7130)
Hind III (5200)
Xho I (6891)
ESTERASE
Xho I (5711)

FIG. 5

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 89/03060

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[5]: C 12 N 15/73, 15/83, 15/82, C 12 P 41/00, C 12 N 15/53, 15/55, 5/10, 7/01, // C 12 N 9/00

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[5]	C 12 N, C 12 P

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
E	WO, A, 89/08145 (BIOSOURCE GENETICS CORP.) 8 September 1989 see the whole document	1-66
X	WO, A, 87/07644 (DIATECH LIMITED) 17 December 1987 see the whole document, especially page 33	1,4-10, 13-19,22-28,31-46, 49-56.
Y	EP, A, 0233656 (GIST-BROCADES N.V.) 26 August 1987 see pages 19-23 cited in the application	1,2,5-11, 14-20,23-29,32-47, 50-56
Y	EP, A, 0194809 (LUBRIZOL GENETICS INC.) 17 September 1986 see the whole document, especially pages 17-23 cited in the application	1,2,4-11, 13-20,22-29,31-47, 49-56

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
30th October 1989	12.12.89
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	T.K. WILLIS

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. US 8903060 SA 30228

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 06/12/89
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A- 8908145	08-09-89	None	
WO-A- 8707644	17-12-87	AU-A- 7489687	11-01-88
		EP-A- 0270611	15-06-88
		GB-A- 2199328	06-07-88
		JP-T- 1500961	06-04-89
EP-A- 0233656	26-08-87	AU-A- 6707286	09-07-87
		JP-A- 63045234	26-02-88
EP-A- 0194809	17-09-86	AU-A- 5437886	11-09-86
		JP-A- 62029984	07-02-87